REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	9 Feb. 1996	Final, 16 Mar 93 - 31 Oct 95

4. TITLE AND SUBTITLE
System Identification for Robust Control Design

5. FUNDING NUMBERS
F49620-93-C-0012

6. AUTHOR(S)
Dr. Robert L. Kosut

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Integrated Systems Inc.
3260 Jay Street
Santa Clara, CA 95054

8. PERFORMING ORGANIZATION REPORT NUMBER
ISI Report 5316-3

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR
Bolling AFB
Washington, D.C.

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
See DODD5230.24

Unlimited

12b. DISTRIBUTION CODE

DISTRIBUTION STATEMENT A
Approved for public release
Distribution Unlimited

13. ABSTRACT (Maximum 200 words)

The long-range goal of this research program is to form a new system identification paradigm that fulfills all the requirements of robust control design, i.e., to produce from finite measured data a model which includes descriptors for both dynamic and disturbance uncertainty.

Several directions were pursued during the reporting period: (1) least-squares related approaches (2) 'windsurfer' (learning) adaptation, and (3) uncertainty model unfalsification. The latter, together with some of the ideas in 'windsurfer' adaptation, marks a significant advance towards attaining the program objectives. The new idea is to replace identification of a single (transfer function) model with unfalsification of a family of uncertainty models and then to design a correponding family of robust controllers to be tested on the actual system. (Both dynamic and noise uncertainty parameters are estimated along with the nominal transfer function).

14. SUBJECT TERMS			15. NUMBER OF PAGES
Identification, robust control			106
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

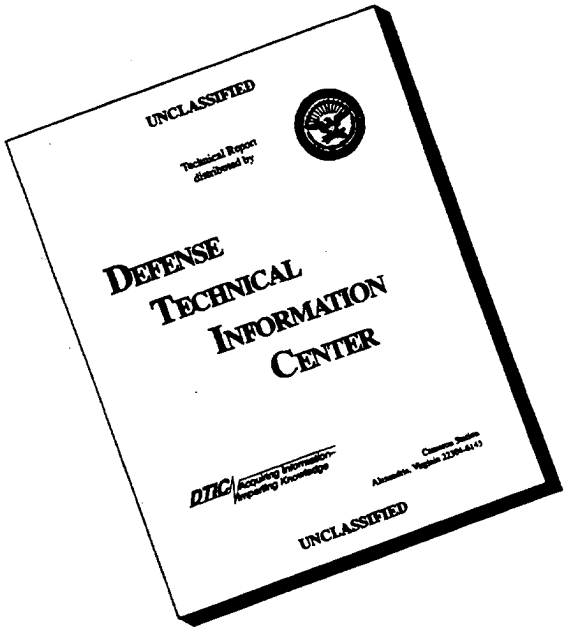
NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std Z39-18
298-102

DTIC QUALITY INSPECTED 1

19960502 091

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

integrated systems

Final Report
16 March 1993 through 31 October 1995

System Identification for Robust Control Design

(Contract No. F49620-93-C-0012)

Approved for public release;
Distribution unlimited

Prepared by:

Dr. Robert L. Kosut
Integrated Systems Inc.
3260 Jay St.
Santa Clara, California 95054
kosut@isi.com,(tel)408-653-0439

Prepared for:

Air Force Office of Scientific Research
Directorate of Mathematical and Computer Sciences
AFOSR/NM, Building 410
110 Duncan Avenue, Suite B115 Bolling Air Force Base
Washington, DC 20332-0001

Attention:

Dr. Marc Jacobs

ISI Report No. 5316-3
9 February 1996

Integrated Systems, Inc.
3260 Jay Street
Santa Clara, CA 95054-3309
Tel: 408.980.1500 Fax: 408.980.0400

Contents

1 Summary of Objectives & Accomplishements

1.1 Objectives

The long-range goal of this research program is to form a new system identification paradigm that fulfills all the requirements of robust control design, *i.e.*, to produce from finite measured data a model which includes descriptors for both dynamic and disturbance uncertainty.

1.2 Accomplishments/New Findings

Several directions were pursued during the reporting period: (1) least-squares related approaches were further investigated, (2) 'windsurfer' adaptation was completed, and (3) uncertainty model unfalsification was developed. The latter, together with some of the ideas in 'windsurfer' adaptation, marks a significant advance towards attaining the program objectives. The new idea is the use of unfalsification in the adaptive control framework. The main ideas are described next.

Inspired by the work of Poolla *et al.*[11], we have shown that **unfalsification** of the standard robust control design uncertainty model is the natural replacement for system identification when the intended use of the model is robust control. For any affine prediction error model, such as an ARX model, the unfalsification step requires solving a set of convex programming problems, specifically LMI problems, of which ordinary least-squares is one member. The result is an **uncertainty tradeoff curve** between model dynamic and disturbance uncertainty. Hence, a family of models are unfalsified from the data record. More specifically:

- For affine models the unfalsification step is an LMI problem with a very specific structure which allows very rapid computations using FFTs. This makes real-time implementation feasible.

- The unfalsification step is downward compatible with existing methods of identification, *i.e.*, one point on the tradeoff curve is identical to classical least-squares prediction error identification.

- Since all uncertainty models along the tradeoff curve are equally unfalsified, or valid, we are led to a natural method of model selection via closed-loop implementation and design iteration. Specifically, to each uncertainty model on the tradeoff curve a robust controller is designed and implemented on the actual system. The process starts with a cautious controller based on a large value of the dynamic uncertainty bound. The bound is gradually decreased, and the controller made more aggressive. Actual closed-loop data is used to evaluate (unfalsify or falsify) the controller. The process continues as long as the actual performance is better than that predicted by the corresponding uncertainty model and worse than that desired. If the desired performance is achieved, then the design is complete. If performance is worse than predicted but not as good as desired, then the uncertainty model has been falsified. This new data can then be used in the next iteration to obtain a new tradeoff curve, *i.e.*, a new family of unfalsified uncertainty models.

The methodology is extensible to many forms of nonlinear systems, but this has not been thoroughly explored.

1.3 Potential Applications

Since system identification enjoys a widespread usage in the engineeering community, it is anticipated that unfalsification methods will be a natural, desireable, and in some cases, an indispensable enhancement to existing software tools. In particular, the aerospace community is probably the largest user of system identification and control design, and as a result a downward compatible replacement which is smoothly compatible with robust control design would be much anticipated.

A very important specific application is adaptive flight control or adaptive reconfigurable control. In this case it is extremely important to have a design model which captures uncertainty, because the reconfigured controller is working in possibly a very hostile environment. In some cases the control objective is to maximize performance of a damaged aircraft. This is a specific case where unfalsification of an uncertainty model would have a significant impact to Air Force missions.

There are also potential applications in civilian technology, in particular, semi-conductor manufacturing, paper mills, and other manufacturing environments which involve high performance requirements of sometimes very uncertain nonlinear systems.

2 Personnel Supported

Prof. Stephen Boyd of Stanford University has been a consultant on the project.

Prof. Anderson of the Australian National University in Canberra, Australia has collaborated with Dr. Kosut on a number of occasions in the past. This year Prof. Anderson has visited Dr. Kosut twice at ISI. Prof. Anderson's is not supported under this contract.

Mr. Wee-Sit Lee, a student of Prof. Anderson's received his Ph. D. this year. Mr. Lee was not directly supported under this contract, but under the guidance of Prof. Anderson and Dr. Kosut, Mr. Lee worked on a method of adaptive control and learning, referred to as "windsurfer adaptation," which has formed the basis for the unfalsification process described above.

3 Publications

Under this contract, Dr. Kosut was an author or co-author of the publications in the list to follow. Copies of all of these publications are included in the Appendix.

1. R.L. Kosut and B.D.O. Anderson, "Statistical analysis of least-squares identification for robust control design: output error case with affine paramtrization," *Proc. 1993 ACC*, San Francisco, CA, June 1993.

2. W.S. Lee, B.D.O. Anderson, R.L. Kosut, and I.M.Y. Mareels, "A new approach to adaptive robust control," *Int.J. Adaptive Control and Signal Processing*, vol. 7, no. 3, pp. 183-212, May-June 1993.

3. H. Aling and R.L. Kosut, "Unbiased least-squares estimates with structure incompatibilities," *Proc. 1993 ECC*, Groningen, The Netherlands, 29 June-2 July, 1993.

4. W.S. Lee, B.D.O. Anderson, I.M.Y. Mareels, and R.L. Kosut, "On some key issues in the windsurfer approach to adaptive control," *Automatica*, Vol. 31, No. 11, pp.1619-1636, 1995. Appeared first in *Proc. SYSID 94*, 10th IFAC Symposium on System Identification, Copenhagen, Denmark, 4-6 July, 1994.

5. M. Massoumnia and R.L. Kosut, "A family of norms for system identification problems," *IEEE Trans. on Automatic Control*, vol. 39, no. 5, May 1994, pp.1027-1031. Appeared first in *Proc. 1993 ACC*, San Francisco, CA, June 1993.

6. R.L. Kosut and B.D.O. Anderson, "Least-squares parameter set estimation for robust control design," *Proc. 1994 ACC*, Baltimore, MD, 29 June-1 July, 1994.

7. R.L. Kosut, "Uncertainty model unfalsification: a new system identification paradigm compatible with robust control design," *Proc. 1995 IEEE CDC*, New Orleans, LA, Dec. 1995.

8. R.L. Kosut, "Iterative adaptive control via uncertainty model unfalsification," invited session, *1996 IFAC World Congress*, San Francisco, CA, June 1996.

4 Interactions/Transitions

4.1 Interactions

1. Dr. Kosut gave a plenary lecture at the *14th Benelux Meeting on Systems and Control* in Houthalen, Belgium, March 29-31, 1995.

2. Dr. Kosut was invited to participate in an NSF/Yale sponsered workshop on "Control using Logic Based Switching," which is another potential applications area for unfalsified uncertainty modeling.

3. Dr. Kosut gave a tutorial lecture on adaptive control at the forthcoming *ISA 95* industrial computing society conference in New Orleans, Oct. 1-6.

4. Dr. Kosut, Prof. Gevers (Louvain University, Belgium), Prof. Bitmead (Autralian National University), and Prof. Smith (UC/Santa Barbara) will give a tutorial workshop at the forthcoming 1996 IFAC World Congress in San Francisco on the interplay between system identification and robust control design.

4.2 Transitions

Since the conceptual framework leading to the unfalsified uncertainty modeling and tradeoff curve has only recently been developed, no transitions have taken place. It is anticipated that this will occur, but first it is crucial to develop reliable and efficient software which can handle typical data sets, which can be quite large. Moreover, for practical use, these tools must be smoothly linked with current robust control design tools.

5 Inventions, or Patent Disclosures

None.

6 Expanded Discussion

6.1 Assessment of State-of-the-Art

Robust control design and system identification have each reached a level of acceptance by the engineering community because of their sound theoretical foundations and the wide availability of software tools based on reliable computational algorithms. However, these two methodologies do not have a common basis, and are at the least, philosophically incompatible.

As described by Ljung [10], traditional transfer function identification postulates an *uncertainty model* consisting of unknown transfer function coefficients and a random exogenous disturbance with an unknown variance. The unknown parameters are to be estimated from a finite input/output data record. In contrast, the standard uncertainty model for robust control design is more realistic in that it includes additional parameters for describing dynamic uncertainty bounds, thereby capturing model limitations.

The focus of this research effort is to develop a quantitative methodology for simultaneously estimating the transfer function parameters, the noise variance, and the dynamic uncertainty bounds from finite-data. The need for this methodology arises in many cases where adaptive control could achieve a high performance capability otherwise unobtainable.

6.2 Unfalsification

The thesis of this research program is that traditional system identification must be modified in order to be compatible with robust control design. The essence of the modification is to replace system identification of the traditional uncertainty model with *validation* of the uncertainty model used for robust control design. In this context validation is defined as follows:

- Given some data, a model is said to be *validated* if and only if it could have produced the data.

As discussed by many philosophers pondering the basis of scientific knowledge, validation of a theory (or model) is perhaps a misnomer, as one can never *prove* that a model will be able to accurately predict the future. However, new data can *falsify* a model, *i.e.*, the model may prove to be incapable of producing the new data. Hence, a more illuminating term is *unfalsification.* Thus, if a data record unfalsifies a model, that model is good enough to explain the current data, but might be *falsified* by a new data record.

Clearly model unfalsification is a feasibility problem – find a member of the model set that is consistent with the data. But *many* postulated models could have produced the data. So what makes one validated model more special than any other? Identification, it seems, poses an optimization problem: pick the model that has the smallest noise variance. Why? It's not the only unfalsified model.

In this past year we have taken a step towards answering this question by establishing a new system identification paradigm which is truly compatible with robust control design. Inspired by the work of Poolla *et al.*[11], we have shown that unfalsification of the standard robust control design uncertainty model is the natural replacement for system identification when the intended use of the model is robust control [5]. For the ARX model (or any affine prediction error model),

the unfalsification step requires solving a set of convex programming problems specifically LMI problems, of which ordinary least-squares is one member. The result is an **uncertainty tradeoff curve** between model uncertainty and disturbance uncertainty. Hence, a family of models are unfalsified from the data record.

Since all uncertainty models along the tradeoff curve are equally unfalsified, or valid, we are led to a natural method of model selection via closed-loop implementation and design iteration. Specifically, to each uncertainty model on the tradeoff curve a robust controller is designed and implemented on the actual system. The process starts with a cautious controller based on a large value of the dynamic uncertainty bound. The bound is gradually decreased, and the controller made more aggressive. Actual closed-loop data is used to evaluate (unfalsify or falsify) the controller. The process continues as long as the actual performance is better than that predicted by the corresponding uncertainty model and worse than that desired. If the desired performance is achieved, then the design is complete. If performance is worse than predicted but not as good as desired, then the uncertainty model has been falsified. This new data can then be used in the next iteration to obtain a new tradeoff curve, *i.e.*, a new family of unfalsified uncertainty models.

Thus, in summary, two changes must be made in the system identification process: (1) postulate a candidate model set which includes descriptors for both disturbance and dynamic uncertainty, and (2) use the data to produce unfalsified models in the candidate model set. There are clearly many theoretical issues to be resolved, e.g., convergence of the iterations, limitations of performance, etc. In addition, for the iteration to be effective, reliable and rapid computational methods need to be developed. Preliminary results in [5] show that the computations involved are very similar to those of system identification.

6.3 Related and Supporting Research

Replacing identification with unfalsification has been investigated and heralded by a number of researchers. The uncertainty tradeoff curve discussed here was first presented in [6], but at that time there was no unfalsification framework upon which to provide a *raison d'etre* for the result. One of the earliest attempts to formulate an unfalsification problem for robust control oriented uncertainty models can be found in the dissertation of Smith and related papers [16, 15]. The approach is based on frequency domain data with a structured block diagonal uncertainty model. In [11], the model validation problem is posed using time-domain data and both necessary and sufficient conditions are obtained. In the work of Doyle *et al.*[1], a large class of unfalsification and identification problems are cast in an implicit linear fractional framework which clearly shows which problems are solvable and which are hard. Iterating and unfaslifying in closed-loop is outlined in the short essay by Dahleh and Doyle [2], and expanded upon in the paper by Livestone *et al.*[8] and dissertation [9]. Direct controller unfalsification is proposed by Safonov *et al.*[17, 18, 12] which uses computational techniques similar to those formulated by Poolla *et al.*[11].

Other approaches have been put forward which also involve iterating on closed-loop data but use traditional identification rather than unfalsification. These are reviewed in the survey paper by Gevers [3]. The fundamental idea is to make the identification and control design criteria identical by iteratively adjusting data and weighting filters, *e.g.*, [13, 7, 20, 19]. In a related approach, referred to as "windsurfer" adaptation [7], the closed-loop bandwidth is increased every iteration, if possible, and in effect, the maximum closed-loop bandwidth is achieved. Many of these ideas can be hopefully made more precise by replacing the identification step with unfalsification.

References

[1] J.C. Doyle, M. Newlin, F. Paganini, and J. Tierno, "Unifying robustness analysis and system ID," *Proc. 33rd CDC*, pp.3667-3672, Lake Buena Vista, FL, Dec. 1994.

[2] M.A. Dahleh and J.C. Doyle, "From data to control", *Proc. Workshop on Modeling of Uncertainty in Control Systems*, Springer-Verlag, 1992.

[3] M. Gevers, "Towards a joint design of identification and control," *Essays on Control: Perspectives in the Theory and its Applications*, editors: H.L. Trentelman and J.C. Willems, pp. 111-151, Birkhauser, Boston, MA, 1993.

[4] M. Gevers and L. Ljung, "Optimal experiment designs with repect to the intended model application," *Automatica*, vol. 22, no. 5, pp. 543-554, 1986.

[5] R.L. Kosut, "Uncertainty model unfalsification: a system identification paradigm compatible with robust control," submitted, 1995 CDC.

[6] R.L. Kosut (1986), "Adaptive Calibration: An Approach to Uncertainty Modeling and On-Line Robust Control Design," *Proc. 25th IEEE CDC,* Athens, Greece, Dec. 1986.

[7] W.S. Lee, B.D.O. Anderson, and R.L. Kosut, "On adaptive robust control and control-relevant system identification", Tech. Report, Dept. of Systems Engr., Australian National University, 1992.

[8] M.M. Livestone, M.A. Dahleh, and J.A. Farrell, "A framework for robust control based model invalidation", *Proc. 1994 ACC*, pp. 3017-3020, Baltimore, MD, June 1994.

[9] M.M. Livestone, *Identification, Robust Adaptation, and Iterative Schemes*, Ph.D. dissertation, MIT Report LIDS-TH-2273, Oct. 1994.

[10] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, 1987.

[11] K. Poolla, P. Khargonnekar, A. Tikku, J. Krause, and K. Nagpal, "A time-domain approach to model validation", *Proc. 1992 ACC*, Chicago, IL, June 1992, and *IEEE Trans. Aut. Contr.*, vol. 39, no. 5, pp. 951-959, May 1994.

[12] M. G. Safonov and T. C. Tsao, "The unfalsified control concept: a direct path from experiment to controller", presented at *Conference on Feedback Control, Nonlinear Systems, and Complexity*, McGill University, Montreal,Canada, May 6-7, 1994.

[13] R.J.P. Schrama, "Accurate identification for control: the necessity for an iterative scheme", "Special Issue on System Identification for Robust Control Design," *IEEE Trans. Aut. Contr.*, vol. 37, no. 7, July, 1992.

[14] R. S. Smith, Editor, *The Modeling of Uncertainty in Control Systems*, Springer-Verlag, 1993.

[15] R.S. Smith and J.C. Doyle, "Model validation – a connection between robust control and identification", "Special Issue on System Identification for Robust Control Design," *IEEE Trans. Aut. Contr.*, vol. 37, no. 7, July, 1992.

[16] R.S. Smith and J.C. Doyle, "Model invalidation – a connection between robust control and identification", *Proc. 1989 ACC*, pp.1435-1440, 1989.

[17] T. C. Tsao and M. G. Safonov, "Set theoretic adaptor systems", *Proc. 1993 ACC*, pp. 3043-3047, San Francisco, CA, June 1993.

[18] T. C. Tsao and M. G. Safonov, "Convex set theoretic adaptor control systems", *Proc. 1993 CDC*, pp. 582-584, San Antonio, TX, Dec 1993.

[19] Y. Yam, D.S. Bayard, R.E. Scheid, "Integrated identification and robust control tuning for large space structures", JPL report, 1989.

[20] Z. Zang, R.R. Bitmead, and M. Gevers, "Iterative model refinement and control robustness enhancement", Tech. Report, Dept. of Systems Engineering, ANU, Nov. 1991.

A Reprints of Publications

This Appendix contains copies of the following publications.

1. R.L. Kosut and B.D.O. Anderson, "Statistical analysis of least-squares identification for robust control design: output error case with affine paramtrization," *Proc. 1993 ACC*, San Francisco, CA, June 1993.

2. W.S. Lee, B.D.O. Anderson, R.L. Kosut, and I.M.Y. Mareels, "A new approach to adaptive robust control," *Int.J. Adaptive Control and Signal Processing*, vol. 7, no. 3, pp. 183-212, May-June 1993.

3. H. Aling and R.L. Kosut, "Unbiased least-squares estimates with structure incompatibilities," *Proc. 1993 ECC*, Groningen, The Netherlands, 29 June-2 July, 1993.

4. W.S. Lee, B.D.O. Anderson, I.M.Y. Mareels, and R.L. Kosut, "On some key issues in the windsurfer approach to adaptive control," *Automatica*, Vol. 31, No. 11, pp.1619-1636, 1995. Appeared first in *Proc. SYSID 94*, 10th IFAC Symposium on System Identification, Copenhagen, Denmark, 4-6 July, 1994.

5. M. Massoumnia and R.L. Kosut, "A family of norms for system identification problems," *IEEE Trans. on Automatic Control*, vol. 39, no. 5, May 1994, pp.1027-1031. Appeared first in *Proc. 1993 ACC*, San Francisco, CA, June 1993.

6. R.L. Kosut and B.D.O. Anderson, "Least-squares parameter set estimation for robust control design," *Proc. 1994 ACC*, Baltimore, MD, 29 June-1 July, 1994.

7. R.L. Kosut, "Uncertainty model unfalsification: a new system identification paradigm compatible with robust control design," *Proc. 1995 IEEE CDC*, New Orleans, LA, Dec. 1995.

8. R.L. Kosut, "Iterative adaptive control via uncertainty model unfalsification," invited session, *1996 IFAC World Congress*, San Francisco, CA, June 1996.

A.1 Statistical analysis of least-squares identification for robust control design: output error case with affine paramtrization

R.L. Kosut and B.D.O. Anderson, *Proc. 1993 ACC*, San Francisco, CA, June 1993.

Statistical Analysis of Least-Squares Identification for Robust Control Design: Output Error Case With Affine Parametrization

Robert L. Kosut*
Integrated Systems, Inc.
3260 Jay St.
Santa Clara, CA 95054, U.S.A.

Brian D. O. Anderson†
Dept. Of Systems Engineering
The Australian National University
Canberra, ACT 2601, Australia

Abstract Precise, finite-data statistical propereties are determined using a least-squares estimator based on an output error model with an affine parameter representation where the true system is of output error form, ***but is not in the model set.*** The purpose of the analysis is to show the effect of unmodeled dynamics on the resulting closed-loop system designed on the basis of the estimated transfer function. This simple problem set-up is prototypical of the interplay between system identification and robust control design.

Introduction

The problem addressed is the following: given a finite collection of sensed sampled input/output data from an unknown system, what level of confidence can be assigned to a feedback controller design or modification.

To make the problem both representative and analytically tractable, the following *a priori qualitative* data is assumed:

(a1) The system which is generating the data is a discrete linear-time-invariant system in *output error* form, *i.e.*,

$$y_t = (Gu)_t + e_t \tag{1}$$

where t is the sampling time, u and y are the sensed input and output sequences, respectively, and e is an unpredictable output disturbance. The operator G is linear-time-invariant with unknown transfer function $G(z)$ and corresponding impulse response sequence g. Thus,

$$(Gu)_t = \sum_{k=1}^{\infty} g_k u_{t-k} \tag{2}$$

(a2) $G(z)$ is stable, *i.e.*, all the poles of $G(z)$ are strictly inside the unit circle. Hence, there exist positive constants $M \geq 1$ and $\rho < 1$ such that

$$|g_k| \leq M\rho^{k-1}, \ \forall k \geq 1 \tag{3}$$

(a3) The unpredictable sequence e is zero-mean gaussian i.i.d. with unknown variance λ_e.

(a4) The input sequence u is deterministic, hence, independent of e.

It is important to emphasize that none of the parameters that appear in the above assumptions are assumed to be known; they are only known to exist. Hence, there is no ***quantitative a priori*** knowledge about M, ρ, or λ_e.

The above qualitative assumptions do, however, impose varying degrees of restrictiveness. Assumption (a1) imposes an LTI structure, which by itself is not necessarily restrictive, however, the output error form is very specific. This latter restriction, together with the gaussian assumption (a3) makes the statistical analysis easier without resorting to a central limit theorem or a law of large numbers. Assumption (a4) implies that the system is operating in open-loop, for otherwise u would have a component which is correlated with e.

For control design it is desireable to obtain an estimate of $G(z)$. It is standard practice to form a parametric model $G(z,\theta)$ and estimate the free parameter θ. Although many parametric forms are possible, *e.g.*, [4], for ease of analysis we choose the following *affine* FIR paramtrization:

$$G(z,\theta) = \sum_{k=1}^{n} \theta_k z^{-k} \tag{4}$$

Thus, the problem is to estimate the first n impulse response coefficients $\{g_1, \cdots, g_n\}$. Although we specialize to the FIR modeling case, all the results apply *mutadis mutandis* to any other affine model of $G(z)$, *e.g.*, Laguerre or Kautz models as described in [5]. The essence of the problem addressed here is, in our opinion, the motivation for the work described in the recent special issue [6] on system identification for robust control design. In comparison with [2], the smoothness parameters M, ρ are not estimated by modeling the tail of the impulse response $\{g_{n+1}, g_{n+2}, \ldots\}$ as a random variable. Our attempt here is to precisely determine the *effect* of the unmodeled dynamics, *i.e.*, the tail of the impulse response, on a least-squares parameter estimator, without any further prior assumptions.

*Supported by AFOSR, Directorate of Mathematical and Computer Sciences, Contract No. F49620-90-C-0064, NSF U.S./Australia Cooperative Research Program, Travel Grant INT-9014152, and the Australian National University as a Visiting Fellow.

†Supported by the Cooperative Rersearch Center for Robust and Adaptive Control

Least-Squares Estimation

In this section we use least-squares on the measured data to estimate the first n impulse response coefficients $\{g_1, g_2, \ldots\}$ in (2). Towards this end, the unknown impulse response parameters $\{g_1, \ldots, g_L\}$ are partitioned into the (finite) parameter vector to be estimated,

$$\alpha = \begin{bmatrix} g_1 \\ \vdots \\ g_n \end{bmatrix} \in \mathbb{R}^n \tag{5}$$

which consists of the first n impulse response coefficients, and the (infinite) parameter vector

$$\beta = \begin{bmatrix} g_{n+1} \\ g_{n+2} \\ \vdots \end{bmatrix} \in \mathbb{R}^\infty \tag{6}$$

which is the remainder of the impulse response. These parameters – the "tail" of the impulse response, $\{g_{n+1}, g_{n+2}, \ldots\}$ – can significantly bias the estimate of the "head," namely, $\{g_1, \ldots, g_n\}$. Statisticians refer to β as a "nuisance" parameter. Note that because G is stable, $\|\beta\|$ is not only finite, but decreases exponentially as n increases. That is, using (3),

$$\|\beta\|^2 = \sum_{k=n+1}^{\infty} g_k^2 \leq \frac{M^2 \rho^{2n}}{1-\rho^2} \tag{7}$$

Using the definition of α and β together with (1) gives,

$$Y = X\alpha + \widetilde{X}\beta + E \tag{8}$$

where

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \in \mathbb{R}^N, \quad E = \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix} \in \mathbb{R}^N \tag{9}$$

$$X = \begin{bmatrix} u_0 & \cdots & u_{1-n} \\ \vdots & & \vdots \\ u_{N-1} & \cdots & u_{N-n} \end{bmatrix} \in \mathbb{R}^{N \times n} \tag{10}$$

$$\widetilde{X} = \begin{bmatrix} u_{-n} & u_{-n-1} & \cdots \\ \vdots & & \vdots \\ u_{N-n-1} & u_{N-n-2} & \cdots \end{bmatrix} \in \mathbb{R}^{N \times \infty} \tag{11}$$

Assuming that $X'X \in \mathbb{R}^{n \times n}$ is non-singular, *i.e.*, u is persistently exciting of order n, the least-squares estimate of α is given by the well known formula:

$$\widehat{\alpha} = \begin{bmatrix} \widehat{g}_1 \\ \vdots \\ \widehat{g}_n \end{bmatrix} = \arg \min_{\theta \in \mathbb{R}^n} \|Y - X\theta\|^2 = (X'X)^{-1}X'Y \tag{12}$$

where $\{\widehat{g}_k \mid k = 1:n\}$ can be thought of as estimates of $\{g_k \mid k = 1:n\}$. We also take the estimate of λ_e, the output error variance, as the sample-variance,

$$\widehat{\lambda}_e = \frac{1}{N}\|Y - X\widehat{\alpha}\|^2 \tag{13}$$

When $\beta = 0$, it is well known that $\widehat{\alpha}$ and $\widehat{\lambda}_e$ are the maximum likeihood estimates of α and λ_e, respectively, *e.g.*, [1]. In our case, $\beta \neq 0$, and its effect on the estimates is the subject of the next section.

Statistical Analysis

In this section we analyze the effect of the nuisance parameter β on the estimates $\widehat{\alpha}$ and $\widehat{\lambda}_e$ of α and λ_e, respectively. We use the standard notation $\mathcal{N}(\mu, \Sigma)$ to denote a gaussian distribution with mean μ and variance Σ. Likewise, $\chi^2(m)$ denotes a chi-squared distribution with m degrees of freedom. Recall that if $q \in \mathbb{R}^m$ is drawn from $\mathcal{N}(0, R)$ with R non-singular, then $q'R^{-1}q \in \chi^2(m)$. We also use $\chi^2(m, \tau)$ to denote a *non-central* chi-squared distribution with m degrees of freedom and non-centrality parameter τ. To fix the definition of the non-centrality parameter, if $q \in \mathbb{R}^m$ is drawn from $\mathcal{N}(\mu, R)$, then $q'R^{-1}q \in \chi^2(m, \tau)$ with $\tau = \mu'R^{-1}\mu$. From [3], we also use: as either[1] m or $\tau \to \infty$, $\chi^2(m, \tau) \to \mathcal{N}(m + \tau, 2(m + 2\tau))$. Hence, $\chi^2(m, 0) = \chi^2(m)$ and as $m \to \infty$, $\chi^2(m) \to \mathcal{N}(m, 2m)$.

It is convenient to define the "covariance" matrices,[2]

$$\Sigma_{11} = \frac{1}{N}X'X \in \mathbb{R}^{n \times n} \tag{14}$$

$$\Sigma_{12} = \frac{1}{N}X'\widetilde{X} \in \mathbb{R}^{n \times \infty} \tag{15}$$

$$\Sigma_{22} = \frac{1}{N}\widetilde{X}'\widetilde{X} \in \mathbb{R}^{\infty \times \infty} \tag{16}$$

Observe that only Σ_{11} can be formed from the data and by assumption is invertible.

The following theorem describes the distributions of the key random variables.

Theorem 1 *Define the parameter error,*

$$\widetilde{\alpha} = \widehat{\alpha} - \alpha \tag{17}$$

and the output error,

$$\widehat{E} = Y - X\widehat{\alpha} \tag{18}$$

Under assumptions (a1)-(a4),

(i) The parameter error $\widetilde{\alpha}$ and the residual $\widehat{E}$ are independent and normally distributed as follows:

$$\widetilde{\alpha} \in \mathcal{N}\left(\Sigma_{11}^{-1}\Sigma_{12}\beta, \ \frac{\lambda_e}{N}\Sigma_{11}^{-1}\right) \tag{19}$$

$$\widehat{E} \in \mathcal{N}\left(\Gamma\widetilde{X}\beta, \ \lambda_e \cdot \Gamma\right) \tag{20}$$

where $\Gamma \in \mathbb{R}^{N \times N}$, given by,

$$\Gamma = I_N - X(X'X)^{-1}X' \tag{21}$$

has rank $N - n$ and is idempotent, i.e., $\Gamma = \Gamma^2$.

[1] It can be shown that this result is also true if both m or $\tau \to \infty$.

[2] Although the matrices Σ_{12}, Σ_{22} are infinite dimensional, they always appear multiplying β. Hence, these terms are bounded because the elements in β decay exponentially.

(ii) $\frac{N\widehat{\lambda}_e}{\lambda_e}$ *and* $\frac{N}{\lambda_e}\widetilde{\alpha}'\Sigma_{11}\widetilde{\alpha}$ *have the following non-central chi-squared distributions:*

$$\frac{N\widehat{\lambda}_e}{\lambda_e} \in \chi^2\left(N-n,\ \frac{N}{\lambda_e}\delta\right) \quad (22)$$

$$\frac{N}{\lambda_e}\widetilde{\alpha}'\Sigma_{11}\widetilde{\alpha} \in \chi^2\left(n,\ \frac{N}{\lambda_e}\gamma\right) \quad (23)$$

where

$$\gamma = \beta'\Sigma_{12}'\Sigma_{11}^{-1}\Sigma_{12}\beta \quad (24)$$

$$\delta = \beta'\Sigma_{22}\beta - \gamma = \beta'\left(\Sigma_{22} - \Sigma_{12}'\Sigma_{11}^{-1}\Sigma_{12}\right)\beta \quad (25)$$

(iii) As $N \to \infty$,

$$\widehat{\lambda}_e \to \mathcal{N}\left(\lambda_e+\delta,\ \frac{2\lambda_e}{N}(\lambda_e+2\delta)\right) \quad (26)$$

$$\widetilde{\alpha}'\Sigma_{11}\widetilde{\alpha} \to \mathcal{N}\left(\frac{n\lambda_e}{N}+\gamma,\ \frac{2\lambda_e}{N}\left(\frac{n\lambda_e}{N}+2\gamma\right)\right) \quad (27)$$

The results in part(i) follow directly from the underlying assumptions and definitions of the variables, and except for the non-zero bias terms, are standard, *e.g.*, [1]. Part (ii) is non-standard, in that the error statistics involve non-central chi-square distributions. These results are obtained by direct appeal to the relation between a normally distributed random variable and the non-central chi-squred statistic as stated in the introduction to this section. The asymptotic results in part(iii) follow from the asymptotic normal approximation to a non-central chisquare distribution as stated in the introduction to this section.

In part (iii) of the theorem, the asymptotic variances decay as $1/N$. Hence, for sufficiently large N, the random variable approaches the mean with high probability. This leads directly to the following:

Approximation 1 *For sufficiently large* N, *the following approximations hold with high probability,*

$$\widehat{\lambda}_e \approx \lambda_e + \delta \quad (28)$$

$$\widetilde{\alpha}'\Sigma_{11}\widetilde{\alpha} \approx \frac{n}{N}\lambda_e + \gamma \quad (29)$$

Observe that for large N, the variance estimate $\widehat{\lambda}_e$ tends to over-estimate the true variance λ_e. In addition, the errors $\widetilde{\alpha}$ and $\widehat{\lambda}_e - \lambda_e$ are driven by the "nuisence" parameter β, *i.e.*, the tail of the impulse response.

A special case of interest is when the input u is white, *i.e.*,

$$\Sigma_{11} = \lambda_u \cdot I_n,\ \Sigma_{12} = 0,\ \Sigma_{22} = \lambda_u \cdot I_\infty \quad (30)$$

Theorem 2 *If* u *is white,* i.e., *(30) holds, then:*

$$\frac{N\widehat{\lambda}_e}{\lambda_e} \in \chi^2\left(N-n,\ \frac{N}{\lambda_e}\lambda_u\|\beta\|^2\right) \quad (31)$$

$$\frac{N}{\lambda_e}\lambda_u\|\widetilde{\alpha}\|^2 \in \chi^2(n) \quad (32)$$

In addition, as $N \to \infty$,

$$\widehat{\lambda}_e \to \mathcal{N}\left(\lambda_e+\lambda_u\|\beta\|^2,\ \frac{2\lambda_e}{N}\left(\lambda_e+2\lambda_u\|\beta\|^2\right)\right) \quad (33)$$

The asymptotic part of the above theorem leads to the following:

Approximation 2 *For sufficiently large* N, *if* u *is white,* i.e., *(30) holds, then with high probability:*

$$\widehat{\lambda}_e \approx \lambda_e + \lambda_u\|\beta\|^2 \quad (34)$$

$$\|\widetilde{\alpha}\|^2 \leq \frac{3n}{N}\frac{\lambda_e}{\lambda_u} \quad (35)$$

Large N and High Probability

When the input is white, "large N" can be taken as,

$$N >> \frac{2(1+2\eta)}{(1+\eta)^2},\ \eta = \frac{\lambda_u\|\beta\|^2}{\lambda_e} \quad (36)$$

where η is the ratio of the energy in the tail to the output error energy. Typical values of N, *e.g.*, 500-1000, will always be well in excess of variations caused by η. Moreover, from central and non-central chi-square tables (*e.g.*, [3]), values of $N \geq 100$ and $n \geq 20$ make the normal approximations very accurate. In consequence, "high probability" is in excess of 99.95% for typical data lengths and model orders. Similar numbers hold for the general case with a non-white input.

Frequency Response Estimation

The results of the previous section can be used to analyze the errors in frequency response estimation. Towards this end, express $G(z)$, the true transfer function as,

$$G(z) = D(z)'\alpha + \widetilde{D}(z)'\beta \quad (37)$$

where

$$D(z) = \begin{bmatrix} z^{-1} \\ \vdots \\ z^{-n} \end{bmatrix},\ \widetilde{D}(z) = \begin{bmatrix} z^{-(n+1)} \\ z^{-(n+2)} \\ \vdots \end{bmatrix} \quad (38)$$

Let $\widehat{G}(z)$ denote the transfer function estimate of $G(z)$ defined as

$$\widehat{G}(z) = D(z)'\widehat{\alpha} \quad (39)$$

where $\widehat{\alpha}$ is the least-squares parameter estimate from (12) of the the first n impulse response coefficients of $G(z)$. Let $\Delta(z)$ denote the transfer function error defined as,

$$\Delta(z) = G(z) - \widehat{G}(z) \quad (40)$$

$$= -D(z)'\widetilde{\alpha} + \widetilde{D}(z)'\beta \quad (41)$$

where

$$D(z)'\widetilde{\alpha} = \sum_{k=1}^{n}(\widehat{g}_k - g_k)z^{-k},\ \widetilde{D}(z)'\beta = \sum_{k=n+1}^{\infty} g_k z^{-k} \quad (42)$$

with $\widetilde{\alpha}$ the parameter error from (17).

From Theorem 1 the following result is obtained.

Theorem 3 *The following results hold at each frequency* ω:

(i) Normal distribution

$$\Delta(e^{j\omega}) \in \mathcal{N}\left(F(e^{j\omega})'\beta,\ \frac{\lambda_e}{N} D(e^{j\omega})^* \Sigma_{11}^{-1} D(e^{j\omega})\right) \quad (43)$$

where

$$F(z)' = \widetilde{D}(z)' - D(z)'\Sigma_{11}^{-1}\Sigma_{12} \quad (44)$$

(ii) Non-central chi-squared distribution

$$\frac{|\Delta(e^{j\omega})|^2}{\frac{\lambda_e}{N} D(e^{j\omega})^* \Sigma_{11}^{-1} D(e^{j\omega})} \in \chi^2\left(1, \epsilon(\omega)\right) \quad (45)$$

with non-centrality parameter,

$$\epsilon(\omega) = \frac{|F(e^{j\omega})'\beta|^2}{\frac{\lambda_e}{N} D(e^{j\omega})^* \Sigma_{11}^{-1} D(e^{j\omega})} \quad (46)$$

(iii) Asymptotic Normality

As $N \to \infty$,

$$\frac{|\Delta(e^{j\omega})|^2}{\frac{\lambda_e}{N} D(e^{j\omega})^* \Sigma_{11}^{-1} D(e^{j\omega})} \to \mathcal{N}\left(1 + \epsilon(\omega),\ 2\left(1 + 2\epsilon(\omega)\right)\right) \quad (47)$$

Part (iii) leads to the following result.

Approximation 3 *For sufficiently large N, the following approximation holds with high probability at each frequency ω:*

$$|\Delta(e^{j\omega})|^2 \approx \frac{\lambda_e}{N} D(e^{j\omega})^* \Sigma_{11}^{-1} D(e^{j\omega}) + |F(e^{j\omega})'\beta|^2 \quad (48)$$

Observe that if u is white (30) then

$$\begin{aligned} D(e^{j\omega})^* \Sigma_{11}^{-1} D(e^{j\omega}) &= D(e^{j\omega})^* \left(\frac{1}{\lambda_u} I_n\right) D(e^{j\omega}) \\ &= \frac{1}{\lambda_u} D(e^{j\omega})^* D(e^{j\omega}) = \frac{n}{\lambda_u} \end{aligned}$$

This leads to the following:

Theorem 4 *If u is white,* i.e., *(30) holds, then at each frequency ω:*

(i) Normal distribution

$$\Delta(e^{j\omega}) \in \mathcal{N}\left(\widetilde{D}(e^{j\omega})'\beta,\ \frac{n}{N}\frac{\lambda_e}{\lambda_u}\right) \quad (49)$$

(ii) Non-central chi-squared distribution

$$\frac{|\Delta(e^{j\omega})|^2}{\left(\frac{n}{N}\frac{\lambda_e}{\lambda_u}\right)} \in \chi^2\left(1, \epsilon(\omega)\right) \quad (50)$$

with non-centrality parameter

$$\epsilon(\omega) = \frac{|\widetilde{D}(e^{j\omega})'\beta|^2}{\left(\frac{n}{N}\frac{\lambda_e}{\lambda_u}\right)} \quad (51)$$

(iii) Asymptotic Normality

As $N \to \infty$,

$$\frac{|\Delta(e^{j\omega})|^2}{\left(\frac{n}{N}\frac{\lambda_e}{\lambda_u}\right)} \to \mathcal{N}\left(1 + \epsilon(\omega),\ 2\left(1 + 2\epsilon(\omega)\right)\right) \quad (52)$$

Part (iii) together with Approximation 2 leads to:

Approximation 4 *If u is white,* i.e., *(30) holds, then for sufficiently large N, the following approximation holds with high probability at each frequency ω:*

$$|\Delta(e^{j\omega})|^2 \approx \frac{n}{N}\frac{\lambda_e}{\lambda_u} + |\widetilde{D}(e^{j\omega})'\beta|^2 \quad (53)$$

Robust Control Analysis

In this section, we use the asymptotic frequency domain bounds to evaluate controller robustness. The goal of control is to reduce the output variance. Consider the LTI feedback controller

$$u = -\widehat{K}y \quad (54)$$

where $\widehat{K}$ stabilizes the "estimated" FIR system

$$y = \widehat{G}u + e,\ \widehat{G}(z) = \sum_{k=1}^{n} g_k z^{-k} \quad (55)$$

Applying the control (54) to the actual system (1) yields the closed-loop system

$$y = \frac{\widehat{T}}{1 + \widehat{Q}\Delta} e,\ u = -\frac{\widehat{Q}}{1 + \widehat{Q}\Delta} e \quad (56)$$

where

$$\widehat{T} = \frac{1}{1 + \widehat{G}\widehat{K}},\ \widehat{Q} = \frac{\widehat{K}}{1 + \widehat{G}\widehat{K}} \quad (57)$$

with Δ the estimation error as defined in (40). Since the nominal system is stable, it follows that Δ, $\widehat{T}$, and $\widehat{Q}$ are stable transfer functions. Hence, the closed-loop system is stable if and only if,

$$|1 + \widehat{Q}(e^{j\omega})\Delta(e^{j\omega})| > 0,\ \forall |\omega| \le \pi \quad (58)$$

If this holds, then the spectrum of y, under closed-loop –*not during identification*– is given by:

$$\Phi_y(\omega) = \left|\frac{\widehat{T}(e^{j\omega})}{1 + \widehat{Q}(e^{j\omega})\Delta(e^{j\omega})}\right|^2 \lambda_e \quad (59)$$

Suppose that u, *during identification*, is white, *i.e.*, (30) holds. To establish stability, observe that a sufficient condition for stability is that,

$$|\widehat{Q}(e^{j\omega})| \cdot |\Delta(e^{j\omega})| < 1,\ \forall |\omega| \le \pi \quad (60)$$

Using the expression for $|\Delta(e^{j\omega})|$ in Approximation 4 and substituting for λ_e from (34). it follows that for large N, the closed-loop system is stable, with high probability, if,

$$|\widehat{Q}(e^{j\omega})|^2 \left[\frac{3n}{N}\left(\frac{\widehat{\lambda}_e}{\lambda_u} - \|\beta\|^2\right) + |\widetilde{D}(e^{j\omega})'\beta|^2\right] < 1, \ \forall|\omega| \leq \pi \tag{61}$$

Hence, using the large N approximations, with high probability, the output spectrum is bounded as follows:

$$\Phi_y(\omega) \leq \frac{|\widehat{T}(e^{j\omega})|^2\left(\widehat{\lambda}_e - \lambda_u\|\beta\|^2\right)}{\left(1 - |\widehat{Q}(e^{j\omega})|\left[\frac{3n}{N}\left(\frac{\widehat{\lambda}_e}{\lambda_u} - \|\beta\|^2\right) + |\widetilde{D}(e^{j\omega})'\beta|^2\right]^{1/2}\right)^2} \tag{62}$$

The only unknown quantity is β. From (34), we also know with high probability that,

$$\lambda_e \approx \widehat{\lambda}_e - \lambda_u\|\beta\|^2$$

Since λ_e must be positive, it follows that

$$\|\beta\|^2 \leq \widehat{\lambda}_e/\lambda_u \tag{63}$$

provides a worst-case upper bound. Observe that this bound is known because $\widehat{\lambda}_e$ is the computed variance estimate and λ_u is selected by the user as the input variance. As a practical matter, it is unlikely that β will achieve this bound. If it did, then the noise variance $\lambda_e \approx 0$, which for large N, will almost never occur.

Using (3), we get

$$|\widetilde{D}(e^{j\omega})'\beta| = |\sum_{k=n+1}^{\infty} g_k e^{-j\omega k}| \leq \frac{M\rho^n}{1-\rho}$$

Hence, for large N, the closed-loop system is stable with high probability if,

$$|\widehat{Q}(e^{j\omega})|^2\left[\frac{3n}{N}\frac{\widehat{\lambda}_e}{\lambda_u} + \frac{M^2\rho^{2n}}{(1-\rho)^2}\right] < 1, \ \forall|\omega| \leq \pi \tag{64}$$

The constants M and ρ are unknown, so in order to evaluate the above robustness condition, either we require *a priori* knowledge or infer the values from the first n impulse response coefficients $\widehat{\alpha}' = [\widehat{g}_1 \cdots \widehat{g}_n]$. That is, define the estimates $\widehat{M}, \widehat{\rho}$ via

$$|\widehat{g}_k| \leq \widehat{M}\widehat{\rho}^{k-1}, \ \forall k \in [1, n] \tag{65}$$

and replace M, ρ with $\widehat{M}, \widehat{\rho}$. This leads to the robustness test:

$$|\widehat{Q}(e^{j\omega})|^2\left[\frac{3n}{N}\frac{\widehat{\lambda}_e}{\lambda_u} + \frac{\widehat{M}^2\widehat{\rho}^{2n}}{(1-\widehat{\rho})^2}\right] < 1, \ \forall|\omega| \leq \pi \tag{66}$$

Now, suppose that the closed-loop system is stable and the above inequality holds. Then the spectrum of y is bounded, with high probability, by: of y and u are given, respectively, by:

$$\Phi_y(\omega) \leq \frac{|\widehat{T}(e^{j\omega})|^2\widehat{\lambda}_e}{\left(1 - |\widehat{Q}(e^{j\omega})|\left[\frac{n}{N}\frac{\widehat{\lambda}_e}{\lambda_u} + \frac{\widehat{M}^2\widehat{\rho}^{2n}}{(1-\widehat{\rho})^2}\right]^{1/2}\right)^2} \tag{67}$$

The above bound gives an indication of the trade between bias and variance as the model order varies – all results being valid for data length $N \geq 500$ with probability in excess of 99.95

Concluding Remarks

Using an output error linear plant, we have shown that with gaussian noise and affine models, there is a very rich structure in the analysis of standard least-squares estimation of the first n impulse response coefficients. The remaining coefficients bias the estimate in a precisely defined way involving non-central chi-squared statistics. These appear to be extremely useful in predicting model error for robust control design from finite data records. Much still remains to be done even for this restricted and analytically tractable case, particularly in finding a means to bound the effect of the bias (the tail of the impulse response) without having to perform additional identification with ever larger parameter orders. This ultimately may involve additional *a priori* quantitative knowlwdge. We feel that this paper indicates a first step towards the more difficult problem of model structures which account for non-white noise, *e.g.*, ARX or ARMAX models.

*

References

[1] G.C. Goodwin and R.L. Payne, *Dynamic System Identification: Experiment Design and data Analysis*, Academic Press, New York, 1977.

[2] G.C. Goodwin, M. Gevers, and B. Ninness, "Quantifying the error in estimated transfer functions with applications to model order selection," *IEEE Trans. on Automatic Control*, vol. 37, no.7, pp. 913-928, July 1992.

[3] N.L. Johnson and S. Kotz, *Continuous Univariate Distributions*, 1970.

[4] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, Inc., 1987.

[5] B. Wahlberg, "System identification using Laguerre models," *IEEE Trans. on Automatic Control*, vol. 36, no.5, pp. 551-562, 1991.

[6] "Special Issue on System Identification for Robust Control Design," edited by R.L. kosut, G.C. Goodwin, and M.P. Polis, *IEEE Trans. on Automatic Control*, vol. 37, no.7, pp. 899-1008, July 1992.

A.2 A new approach to adaptive robust control

W.S. Lee, B.D.O. Anderson, R.L. Kosut, and I.M.Y. Mareels, *Int.J. Adaptive Control and Signal Processing*, vol. 7, no. 3, pp. 183-212, May-June 1993.

INTERNATIONAL JOURNAL OF ADAPTIVE CONTROL AND SIGNAL PROCESSING, VOL. 7, 183–211 (1993)

A NEW APPROACH TO ADAPTIVE ROBUST CONTROL

WEE SIT LEE AND BRIAN D. O. ANDERSON
Department of Systems Engineering, The Australian National University, Canberra, ACT 0200, Australia

ROBERT L. KOSUT
Integrated Systems, Inc., 3260 Jay St., Santa Clara, CA 95054, U.S.A.

AND

IVEN M. Y. MAREELS
Department of Systems Engineering, The Australian National University, Canberra, ACT 0200, Australia

SUMMARY

A new approach is given for the design of adaptive robust control in the frequency domain. Starting with an initial model of a stable plant and a robust stabilizing controller, the new (windsurfer) approach allows the bandwidth of the closed-loop system to be increased progressively through an iterative control-relevant system identification and control design procedure. The method deals with both undermodelling and measurement noise issues. Encouraging results are obtained in the simulations that illustrate the new idea.

KEY WORDS Adaptive control Robust control Internal model control System identification

1. INTRODUCTION

It has long been understood that a key problem in control system design is to handle the *uncertainties* associated with the plant.[1] Two main techniques for the analysis and design of systems with significant uncertainties are adaptive control[2] and robust control.[3,4]

In the traditional approach to analysis and design of an adaptive control system[2] it is assumed that the unknown plant can be represented by a model in which everything is known except for the values of a finite number of parameters. Once the parameters are estimated (and even during the estimation process), the principle of certainty equivalence is normally invoked to update the controller. Normally the unstructured uncertainties of the model are *ignored* in this approach. Therefore it is not surprising, as pointed out in Reference 5 that these adaptive controllers are often not robust. Further, the extensions of the traditional approach to adaptive control which purportedly cope with unstructured (and other) uncertainties involve conditions which are often hard to apply or to grasp intuitively (see e.g. References 6–8). A further problem with the traditional approach is that extreme transient excursions are possible even when global convergence and asymptotic performance are guaranteed.[9]

To be more specific, we consider an adaptive control system as shown in Figure 1, where G is the unknown transfer function of the plant. The time axis is divided into intervals such that

This paper was recommended for publication by editor I. Kanellakopoulos

0890–6327/93/030183–29$19.50 *Received February 1993*
© 1993 by John Wiley & Sons, Ltd.

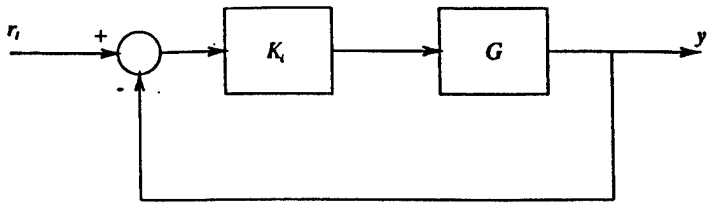


Figure 1. Adaptive control system

during the ith interval the control input applied to the plant is obtained from K_i, where K_i is the transfer function of the controller designed using the model G_{i-1} obtained at the end of the $(i-1)$th time interval.

In an adaptive control problem the ulterior objective for finding G_i, an estimate of G updated from G_{i-1}, is to design a controller K_{i+1} better than K_i such that certain control objectives are improved. For example, if T^d represents the desired complementary sensitivity function, then we may like to have

$$\left\| \frac{GK_i}{1+GK_i} - T^d \right\|_\infty \leqslant \left\| \frac{GK_{i-1}}{1+GK_{i-1}} - T^d \right\|_\infty \quad \forall i$$

Implicitly, this means we would like to minimize

$$\left\| \frac{GK_i}{1+GK_i} - T^d \right\|_\infty \quad \forall i$$

Since G, the transfer function of the plant, is unknown, we could only base our design of K_i on G_{i-1} such that

$$K_i = \arg \min_\gamma \left\| \frac{G_{i-1}\gamma}{1+G_{i-1}\gamma} - T^d \right\|_\infty \quad \forall i$$

Note that, as usual, we have invoked the principle of certainty equivalence. However, it is important to realize that

$$\left\| \frac{GK_i}{1+GK_i} - T^d \right\|_\infty$$

is not necessarily small even though

$$\left\| \frac{G_{i-1}K_i}{1+G_{i-1}K_i} - T^d \right\|_\infty$$

is minimum. This partly explains why traditional adaptive control systems, which almost invariably invoke the principle of certainty equivalence, often have unsatisfactory robustness properties.

In the robust control approach[3,4] a controller is designed on the basis of a nominal model of the plant with the associated parametric and unstructured model uncertainties explicitly taken into account. Therefore stability robustness is guaranteed and performance robustness is achieved sometimes. The *weakness* of this approach is that it considers only the *a priori* information on the model and *neglects* the fact that some characteristics of the plant could be learnt while it is being controlled. Therefore the robust control approach tends to result in a conservative design in terms of performance. It is likely that *a posteriori* knowledge about the plant could be used to reduce the conservatism in a robust control design.

In this paper we present a new approach for the design of adaptive robust control for a stable plant and the related control-relevant system identification. A preliminary study of the problem under noiseless conditions is reported in Reference 10. The main differences of this paper from Reference 10 are that, instead of updating the models through finding rational function approximations (in the H_∞ sense) of certain plant parametrization and employing proper but non-strictly proper controllers, here we perform system identification using input–output data and employ strictly proper controllers. It was also shown in References 11 and 12 that an iterative approach for model refinement and control robustness enhancement can be developed in the context of an H_2 control problem.

In Section 2 we describe the windsurfer approach[13] to adaptive control. In Section 3 we apply the windsurfer approach to an adaptive model-matching problem and formulate the related control-relevant system identification criterion. The relevance of the system identification criterion to the adaptive model-matching problem is explored further in Section 4. In Section 5 we apply Hansen's approach[14–16] to recast the closed-loop system identification problem into an open-loop system identification problem. We also show that with appropriate filtering, the criterion used in the open-loop system identification is highly relevant to the windsurfer approach to adaptive model matching. In Section 6 we develop the relation between the approximate identification and the internal model control (IMC)[4] method of controller design. We present the simulation results in Section 7. In Section 8 we conclude the paper and attempt to give some reasons for the success of the method.

2. THE WINDSURFER APPROACH TO ADAPTIVE CONTROL

By considering how humans learn windsurfing, Anderson and Kosut[13] have made the following observations.

1. The human first learns to control over a limited bandwidth and learning pushes out the bandwidth over which an accurate model of the plant is known.
2. The human first implements a low-gain controller and learning allows the loop to be tightened.

On the basis of these observations, an adaptive robust control design philosophy, *the windsurfer approach*, is proposed in Reference 13. It recognizes that the plant characteristics can differ greatly from the estimated model at any one time, particularly during the initial learning stage. In the new design approach a low-gain controller will first be implemented and the control bandwidth will be small. On the basis of learning a frequency domain description of the plant *in closed loop*, with the learning process progressively increasing the bandwidth over which the plant is accurately known, the controller gain can be increased appropriately over an increasing frequency band. For details see Reference 13. Importantly, in the method suggested, the necessary closed-loop system identification task is transformed into an open-loop system identification problem through the use of coprime fractional representations as discussed in References 14–16.

It was shown recently[17] that the best model for control design cannot be derived from open-loop experiments alone. The controller to be implemented should be taken into account by the system identification experiments. However, this controller is not yet available, since its determination rests on the results of the system identification to be carried out. Hence a general solution to the combination of system identification and control design is necessarily iterative. Although the emphasis of Reference 17 is on the problem of modelling for control purposes, its approach is very similar to that of Reference 13.

In the next section we would like to illustrate the windsurfer approach by considering a model-matching problem in the context of adaptive control.

3. ADAPTIVE MODEL MATCHING

Let G be the unknown transfer function of the plant and let T^d represent a desired complementary sensitivity function. We wish to achieve, through iterative system identification and control design, the minimization of the cost function

$$\left\| \frac{GK}{1+GK} - T^d \right\|_\infty$$

where K is the transfer function of a controller to be designed.

We begin by designing a controller $K_{1,0}$ to stabilize a known initial model G_0, which may be obtained from an open-loop system identification exercise. If $K_{1,0}$ also stabilizes the unknown transfer function G, then we say that $K_{1,0}$ robustly stabilizes G_0. Notice that we use $K_{j,i}$ to denote the jth controller designed using the ith model which has a transfer function G_i. In general, we attach the subscript j,i to a transfer function to denote that it is either specified or derived on the basis of the ith model for the plant at the jth iteration of control design. Since G_0 may involve significant uncertainties, the resulting controller $K_{1,0}$ may not be able to achieve a small value for

$$\left\| \frac{G_0K_{1,0}}{1+G_0K_{1,0}} - T^d \right\|_\infty$$

while robustly stabilizing G_0. In general, we need to consider how to handle the question of securing robust stabilization of G_i by $K_{j,i}$. This is bound up with the question of selection of T^d. It is in fact to be expected that a sequence of T^d will be selected in such a way that the end control objective can be approached in stages. We shall therefore proceed as follows.

In association with each of the models G_i, a sequence of controllers $K_{j,i}$ is designed such that

$$K_{j,i} = \arg\min_{\gamma} \left\| \frac{G_i\gamma}{1+G_i\gamma} - T^d_{ji} \right\|_\infty \quad \forall j \tag{1}$$

where the sequence of functions $T^d_{j,i}$ is specified with $T^d_{j+1,i}$ normally of wider bandwidth than $T^d_{j,i}$. It is also necessary that $T^d_{1,i}$ results in a controller $K_{1,i}$ that robustly stabilizes G_i. A stage will be reached (say when $j = N$) where the bandwidth of the nominal closed-loop transfer function,

$$\bar{T}_{N,i} = \frac{G_iK_{N,i}}{1+G_iK_{N,i}} \tag{2}$$

cannot be increased further without causing the effects of model uncertainties in G_i to be too significant. This occurs when the value of

$$\| T_{N,i} - \bar{T}_{N,i} \|_\infty$$

is no longer small, where

$$T_{N,i} = \frac{GK_{N,i}}{1+GK_{N,i}} \tag{3}$$

is the actual closed-loop transfer function of the system.

At this stage it is necessary to improve the accuracy of the model in such a way that is

relevant to the control objective. This means that we should try to find an updated model G_{i+1} such that

$$G_{i+1} = \arg \min_{\theta} \left\| \frac{GK_{N,i}}{1 + GK_{N,i}} - \frac{\theta K_{N,i}}{1 + \theta K_{N,i}} \right\|_{\infty} \tag{4}$$

Once G_{i+1} is found, we can continue to increase the closed-loop bandwidth by repeating the procedure described for G_i previously. However, G_{i+1} should be used instead of G_i and we specify a new sequence of functions $T^{d}_{j,i+1}$ with $T^{d}_{1,i+1} = T^{d}_{N,i}$. The iterative process is continued until the end control objective is achieved or it is prematurely terminated because of one or more of the following constraints:

(1) fundamental performance limitations due to right-half-plane poles and zeros of the plant and/or models[18]
(2) unstable model is obtained (This is a consequence of our simplified control design method. Appropriate extensions of the control design method[4] allow us to deal with this restriction.)
(3) finite control energy
(4) no further improvements in the identified model can be made for a reasonably large set of input–output measurements.

4. CONTROL-RELEVANT SYSTEM IDENTIFICATION

It should be noted that the system identification criterion formulated in Section 3,

$$G_{i+1} = \arg \min_{\theta} \left\| \frac{GK_{N,i}}{1 + GK_{N,i}} - \frac{\theta K_{N,i}}{1 + \theta K_{N,i}} \right\|_{\infty} \tag{5}$$

would be the formulation of a standard rational function approximation problem provided that G were known. However, as opposed to an approximation problem, we are here dealing with a system identification problem where G is an unknown transfer function and only a finite number of (possibly noisy) input–output measurements are available. Despite this apparent difference, we must emphasize that equation (5) is exactly the dual of the criterion developed by Anderson and Liu[19] in the controller reduction problem based on closed-loop transfer function consideration, where their plant and reduced-order controller are replaced by our controller and estimated model respectively. We can therefore draw a similar conclusion that in our system identification problem there is a reduced weighting placed on the range of frequencies where the loop gain is high. This is very appealing since it agrees with the well-known fact that, for a stable closed-loop system, model errors are more tolerable in the range of frequencies where the loop gain is allowed to be large. More importantly, as we shall explain below, this system identification criterion will enable us to find a new model which allows us to design a closed-loop system with a larger bandwidth than what the original model would allow.

If we rewrite equation (5) in the form

$$G_{i+1} = \arg \min_{\theta} \left\| \left(\frac{1}{1 + GK_{N,i}} \right) \left(\frac{\theta K_{N,i}}{1 + \theta K_{N,i}} \right) \left(\frac{G - \theta}{\theta} \right) \right\|_{\infty} \tag{6}$$

we see immediately that it is the product of the actual sensitivity function

$$\frac{1}{1 + GK_{N,i}}$$

and the nominal complementary-sensitivity-function-weighted multiplicative model error

$$\left(\frac{G_{i+1}K_{N,i}}{1+G_{i+1}K_{N,i}}\right)\left(\frac{G-G_{i+1}}{G_{i+1}}\right)$$

It appears that the frequency-weighting function in equation (6), which involves the unknown actual sensitivity function, cannot be implemented in the system identification procedure. However, we shall show in Section 5 that by recasting the closed-loop system identification into an open-loop system identification problem, we also obtain a system identification criterion which is equivalent to equation (6) but involves only a known frequency-weighting function. Therefore, for the purpose of understanding the effects of the system identification criterion on the identified model, we can treat the frequency-weighting function in equation (6) as a known quantity.

Recall that, as described in Section 3, we were using the model G_i to design a sequence of controllers $K_{j,i}$, with increasing gain over an increasing range of frequencies, such that the closed-loop system has an increasing bandwidth. At the stage where $j = N$, the gain of the controller $K_{N,i}$ has become so large that the high-frequency model uncertainties associated with G_i are no longer insignificant. Any attempt to increase the closed-loop bandwidth further will cause the magnitude of the nominal complementary-sensitivity-function-weighted multiplicative model error

$$\left(\frac{G_iK_{N,i}}{1+G_iK_{N,i}}\right)\left(\frac{G-G_i}{G_i}\right)$$

to become too large at certain frequencies, such that the system may lose performance robustness or even stability robustness. Therefore the gain of the controller $K_{N,i}$ will be limited and the actual sensitivity function

$$\frac{1}{1+GK_{N,i}}$$

will be large beyond the existing limited closed-loop bandwidth. From equation (6) we notice that it is exactly in this range of frequencies, where the actual sensitivity function has large magnitude, that our system identification criterion will penalize the nominal complementary-sensitivity-function-weighted multiplicative model uncertainties of the new model G_{i+1}. We could therefore expect G_{i+1} to have smaller model uncertainties, as compared with G_i, near and beyond the edge of the closed-loop bandwidth that can be achieved with G_i. This will allow us to design controllers $K_{j,i+1}$ that lead to larger closed-loop bandwidth than was possible with G_i. Hence we can say that criterion (6) is control-relevant.

To make concrete the above discussions, we consider an example where the plant has a transfer function

$$G = \frac{9}{s^2 + 0{\cdot}06s + 9} \tag{7}$$

and an initial model with the transfer function

$$G_0 = \frac{1}{s+1} \tag{8}$$

If we employ a controller which has a transfer function of the form

$$K = \frac{\lambda^2(s+1)}{s(s+2\lambda)} \tag{9}$$

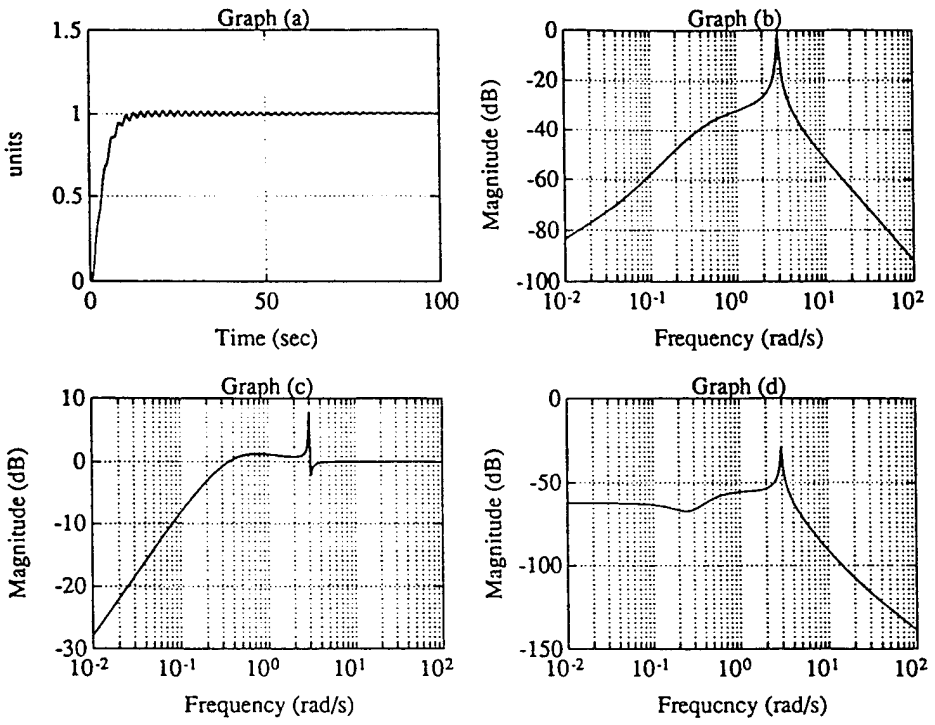


Figure 2. Frequency-weighted multiplicative model errors

then it can be shown that the nominal closed-loop transfer function is given by

$$\bar{T} = \frac{\lambda^2}{(s+\lambda)^2} \tag{10}$$

Therefore λ is the nominal -6 dB bandwidth of the closed-loop system. As λ is increased to $0{\cdot}5$ rad s^{-1}, the actual closed-loop unit-step response, as shown in graph (a) of Figure 2, has excessive oscillations because the model uncertainties associated with G_0 are no longer insignificant. This is apparent in graph (b) of Figure 2, which shows the magnitude of the nominal complementary-sensitivity-function-weighted multiplicative model error of G_0. We have also shown the actual sensitivity function in graph (c) of Figure 2, which indicates that, if it is incorporated into the system identification criterion (for which the procedure is described later in Sections 5 and 6), the nominal complementary-sensitivity-function-weighted multiplicative error in the new model G_1 will be penalized in the range of frequencies near and beyond the existing closed-loop bandwidth of $0{\cdot}5$ rad s^{-1}. It can be seen from graph (d) of Figure 2, which shows the nominal complementary-sensitivity-function-weighted multiplicative error of the new model G_1, that this is indeed the case. Therefore the new model G_1 will allow us to increase the closed-loop bandwidth beyond $0{\cdot}5$ rad s^{-1}.

5. CLOSED-LOOP SYSTEM IDENTIFICATION

We first review a method for closed-loop system identification developed by Hansen and co-workers.[14–16] Subsequently, in Theorem 2, we demonstrate that with appropriate signal filtering, Hansen's method provides a suitable framework to carry out the *control-relevant*

system identification formulated in Section 3. For the sake of expository simplicity we shall consider only scalar plants. We begin with the following theorem.[20]

Theorem 1

If $K = X/Y$ is a controller, where X and Y are stable proper transfer functions, and if N and D are stable proper transfer functions that satisfy the Bezout identity

$$NX + DY = 1$$

then the set of all plants stabilized by the controller K is precisely the set of elements in

$$\mathcal{G} = \left\{ \frac{N + RY}{D - RX} : R \text{ is a stable proper transfer function} \right\}$$

Consider the feedback system shown in Figure 3, where y and u are the measured output and the control input respectively, e is an unpredictable white disturbance and r_1 and r_2 are user-applied inputs. It is assumed that $K_{j,i}$ is a known stabilizing controller, G is inexactly known and possibly unstable and, as is standard,[21] H is imperfectly known, stable and inversely stable. The system identification problem is to obtain improved estimates of G and H from a finite interval of measured and known data $\{y, u, r_1, r_2 : 0 \leqslant t \leqslant T\}$.

Following Hansen,[16] we introduce the stable proper transfer functions $X_{j,i}$, $Y_{j,i}$, $N_{j,i}$ and $D_{j,i}$ which satisfy

$$K_{j,i} = \frac{X_{j,i}}{Y_{j,i}}, \qquad G_i = \frac{N_i}{D_i}, \qquad N_i X_{j,i} + D_i Y_{j,i} = 1$$

The interpretation is that G_i is a known but imperfect model of the plant which is also stabilized by $K_{j,i}$. Applying Theorem 1 as shown in References 15 and 16, there exist stable proper transfer functions $R_{j,i}$ and $S_{j,i}$, with $S_{j,i}$ also inversely stable, such that

$$G = \frac{N_i + R_{j,i} Y_{j,i}}{D_i - R_{j,i} X_{j,i}} \tag{11}$$

$$H = \frac{S_{j,i}}{D_i - R_{j,i} X_{j,i}} \tag{12}$$

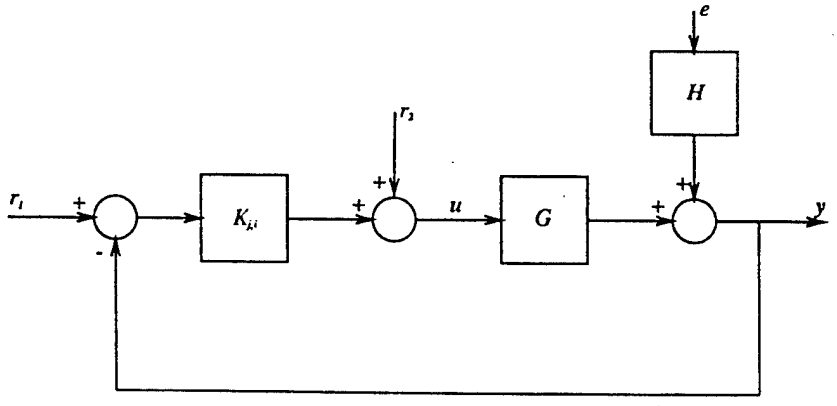


Figure 3. Closed-loop system

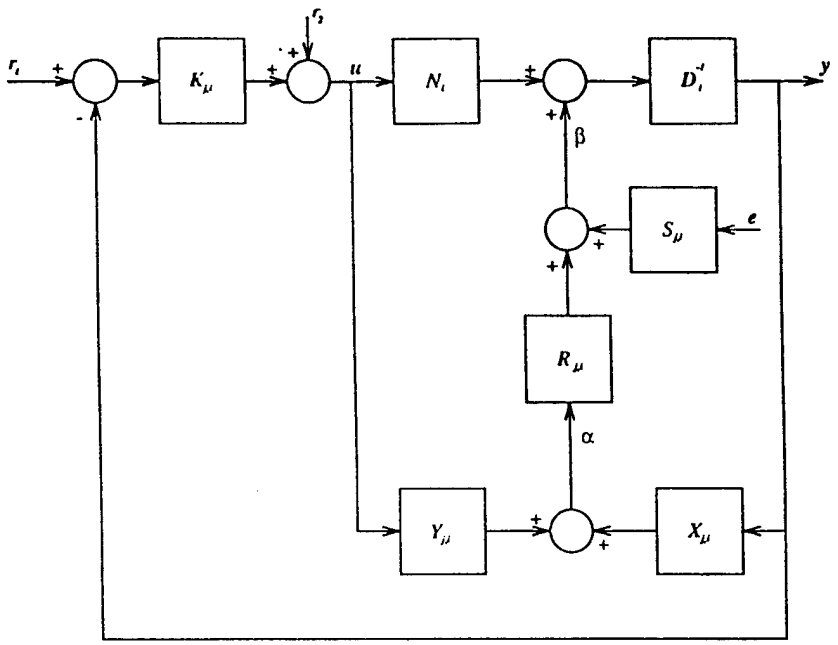


Figure 4. Closed-loop system identification

where $R_{j,i}$ denotes the parametrization of G using the ith model and its associated jth controller $K_{j,i}$.

As a result, system identification of G and H in closed loop is equivalent to system identification of the stable proper transfer functions $R_{j,i}$ and $S_{j,i}$. Using equations (11) and (12), we can represent the feedback system as shown in Figure 4.

From Figure 4 we can write

$$\beta = R_{j,i}\alpha + S_{j,i}e \tag{13}$$

where

$$\alpha = X_{j,i}y + Y_{j,i}u \tag{14}$$

$$\beta = D_i y - N_i u \tag{15}$$

However, since

$$u = K_{j,i}(r_i - y) + r_2$$

and

$$K_{j,i} = \frac{X_{j,i}}{Y_{j,i}}$$

equation (14) can be rewritten as

$$\alpha = X_{j,i}r_1 + Y_{j,i}r_2 \tag{16}$$

It is important to observe from equations (13), (15) and (16) that α depends only on the applied signals r_1 and r_2 operated on by known stable proper transfer functions $X_{j,i}$ and $Y_{j,i}$ respectively, while β depends on the measured signals y and u operated by known stable

proper transfer functions D_i and N_i respectively. Moreover, α is independent of the transfer functions G and H and the disturbance e. Hence the system identification of G and H in closed loop has been recast into the system identification of $R_{j,i}$ and $S_{j,i}$ in open loop.

We shall next state a result which is highly relevant to the system identification step of the windsurfer approach to adaptive control.

Theorem 2

Let the controller $K_{j,i}$ stabilize the plant G and the model

$$G_i = \frac{N_i}{D_i}$$

where N_i and D_i are stable proper transfer functions, and let

$$K_{j,i} = \frac{X_{j,i}}{Y_{j,i}}$$

where $X_{j,i}$ and $Y_{j,i}$ are stable proper transfer functions satisfying the Bezout identity

$$N_i X_{j,i} + D_i Y_{j,i} = 1$$

Let G_{i+1} be another model of G, also stabilized by $K_{j,i}$ and therefore having a description

$$G_{i+1} = \frac{N_i + r_{j,i} Y_{j,i}}{D_i - r_{j,i} X_{j,i}} \tag{17}$$

where $r_{j,i}$ is a stable proper transfer function. Also define the filtered output error

$$\xi = Y_{j,i}(\beta - r_{j,i}\alpha) \tag{18}$$

where, with $r_2 = 0$,

$$\alpha = X_{j,i} r_1, \qquad \beta = D_i y - N_i u$$

$$r_1 = \text{reference signal}, \qquad y = \text{plant output}, \qquad u = \text{control input}$$

Thus ξ is an error arising in the (open-loop) identification of $R_{j,i}$ through an estimate $r_{j,i}$. Then the filtered output error can be expressed as

$$\xi = \left(\frac{GK_{j,i}}{1 + GK_{j,i}} - \frac{G_{i+1}K_{j,i}}{1 + G_{i+1}K_{j,i}} \right) r_1 + \frac{1}{1 + GK_{j,i}} He$$

Proof. See Appendix I. □

Remark

Notice that in Theorem 2 it is necessary that $K_{j,i}$ stabilizes G when the system identification procedure is carried out. This can be ensured by increasing the closed-loop bandwidth smoothly and cautiously in the controller design stages (to be described in Section 6). We would always detect a gradual degradation of performance robustness (while stability is still being maintained and the system identification procedure is being carried out) before the closed-loop system lost stability.

Suppose that the value of

$$\left\| \frac{GK_{j,i}}{1+GK_{j,i}} - \frac{G_iK_{j,i}}{1+G_iK_{j,i}} \right\|_\infty \tag{19}$$

has become large. As it was described in Section 3, we want a new identification of G via G_{i+1} for which

$$\left\| \frac{GK_{j,i}}{1+GK_{j,i}} - \frac{G_{i+1}K_{j,i}}{1+G_{i+1}K_{j,i}} \right\|_\infty \tag{20}$$

is small. We are going to use the $r_{j,i}$-parametrization of G_{i+1}. By substituting equations (11) and (17) into expression (20) and noting that

$$K_{j,i} = \frac{X_{j,i}}{Y_{j,i}}$$

we can, after simplification, conclude that

$$\left\| \frac{GK_{j,i}}{1+GK_{j,i}} - \frac{G_{i+1}K_{j,i}}{1+G_{i+1}K_{j,i}} \right\|_\infty = \| Y_{j,i}X_{j,i}(R_{j,i} - r_{j,i}) \|_\infty \tag{21}$$

should be small.

By using equations (13), (18) and (21), we immediately see that (for the system identification procedure relevant to the windsurfer approach to adaptive control) the appropriate signal model is

$$\beta_1 = R_{j,i}\alpha_1 + v \tag{22}$$

where

$$\beta_1 = Y_{j,i}\beta \tag{23}$$

$$\alpha_1 = Y_{j,i}\alpha \tag{24}$$

and v is the term related to the disturbance e.

Remarks

(i) Note that

$$T_{j,i} = \frac{GK_{j,i}}{1+GK_{j,i}}$$

is the actual closed-loop transfer function of the system and

$$\bar{T}_{j,i} = \frac{G_iK_{j,i}}{1+G_iK_{j,i}}$$

is the nominal closed-loop transfer function of the system. Therefore, using similar substitutions to these that resulted in equation (21), we can obtain

$$T_{j,i} - \bar{T}_{j,i} = Y_{j,i}X_{j,i}(R_{j,i} - \bar{R}_{j,i}) \tag{25}$$

However, since

$$\bar{R}_{j,i} \equiv 0 \quad \forall j, \forall i$$

we therefore have

$$T_{j,i} - \bar{T}_{j,i} = Y_{j,i}X_{j,i}R_{j,i} \tag{26}$$

By comparing the argument of the H_∞-norm given in expression (19) with the left-hand side of equation (26), we see immediately that when the value of

$$\left\| \frac{GK_{j,i}}{1+GK_{j,i}} - \frac{G_iK_{j,i}}{1+G_iK_{j,i}} \right\|_\infty$$

has become large, i.e. when the closed-loop properties of the actual system ($T_{j,i}$) are significantly different from the closed-loop properties of the nominal system ($\bar{T}_{j,i}$), the value of

$$\| Y_{j,i}X_{j,i}R_{j,i} \|_\infty$$

will be large.

(ii) From the signals defined in Theorem 2, we observed that $R_{j,i}$, the transfer function to be identified, is excited by the signal α, where

$$\alpha = X_{j,i}r_1, \qquad X_{j,i} = \frac{K_{j,i}}{1+G_iK_{j,i}}$$

Since the nominal closed-loop transfer function of the system is

$$\bar{T}_{j,i} = \frac{G_iK_{j,i}}{1+G_iK_{j,i}}$$

we can write

$$X_{j,i} = \frac{\bar{T}_{j,i}}{G_i}$$

Therefore $X_{j,i}$ will have large magnitude when we try to push the nominal closed-loop bandwidth beyond the nominal open-loop bandwidth. Since a model usually has its uncertainties become significant for frequencies beyond its bandwidth, from Figure 5 we see that if the spectrum of r_1 is white, we automatically get the right weighting for the input to $R_{j,i}$ for the system identification scheme.

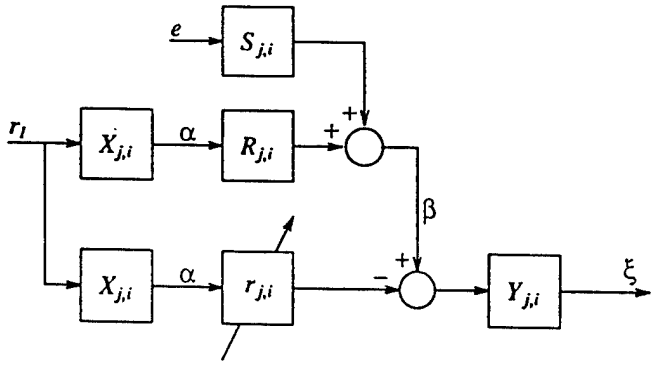


Figure 5. Excitation of $R_{j,i}$

(iii) It is shown in Theorem 2 that the effect of e on ξ is given by

$$\frac{He}{1+GK_{j,i}}$$

Notice that this is the effect of e on y attenuated by the sensitivity function of the actual closed-loop system.

6. APPROXIMATE IDENTIFICATION OF THE $R_{j,i}$ TRANSFER FUNCTION FOR IMC CONTROLLER DESIGN

In Section 5 we have shown that the closed-loop system identification of the plant transfer function G can be reformulated into an open-loop system identification of the stable proper transfer function $R_{j,i}$ that parametrized the transfer function G via the equation

$$G=\frac{N_i+R_{j,i}Y_{j,i}}{D_i-R_{j,i}X_{j,i}}$$

In this and the following sections we shall, for simplicity, study the case where the plant transfer function is stable strictly proper and has no zeros on the imaginary axis of the s-plane and where the IMC method[4] is used to design the controller $K_{j,i}$. We shall also assume that all estimates G_i of the plant are stable strictly proper transfer functions.

If the model

$$G_i=\frac{N_i}{D_i}$$

is stable, we can let $N_i=G_i$ and $D_i=1$ so that

$$G=G_i+\frac{R_{j,i}}{1-R_{j,i}Q_{j,i}} \tag{27}$$

where $Q_{j,i}$ is a stable strictly proper transfer function that parametrized the strictly proper controller

$$K_{j,i}=\frac{X_{j,i}}{Y_{j,i}}$$

and

$$Q_{j,i}\overset{\text{def}}{=}\frac{K_{j,i}}{1+G_iK_{j,i}} \tag{28}$$

The reason for requiring $Q_{j,i}$ and hence $K_{j,i}$ to be strictly proper is that this is a necessary condition for the system to be robust in the presence of high-frequency parasitic or singular perturbation.[22] We also have

$$X_{j,i}=Q_{j,i} \qquad Y_{j,i}=1-Q_{j,i}G_i$$

Since the parametrization of G by $R_{j,i}$ depends intimately on $Q_{j,i}$, we shall briefly explain how $Q_{j,i}$ is obtained in the design of the controller $K_{j,i}$. We will use the notations n_H and d_H to denote the numerator polynomial and the denominator polynomial respectively of a rational transfer function H.

Given a stable model

$$G_i = \frac{n_{G_i}}{d_{G_i}}$$

where d_{G_i} has no zeros in the closed right-half s-plane, if n_{G_i} has no zeros on the imaginary axis of the s-plane, we can write

$$G_i = \frac{\tilde{n}_{G_i} \Pi_i (z_i - s)}{d_{G_i}}$$

where all z_i have positive real parts and $\tilde{n}_{G_i}$ has no zeros in the closed right-half s-plane. By writing G_i as

$$G_i = [G_i]_m [G_i]_a$$

where

$$[G_i]_m = \frac{\tilde{n}_{G_i} \Pi_i (z_i^* + s)}{d_{G_i}}$$

(z_i^* is the complex conjugate of z_i)

$$[G_i]_a = \frac{\Pi_i (z_i - s)}{\Pi_i (z_i^* + s)}$$

we have factored G_i as a product of its minimum phase factor $[G_i]_m$ and the associated all-pass factor $[G_i]_a$. We can find the controller

$$K_{j,i} = \arg \min_{\gamma} \left\| \frac{G_i \gamma}{1 + G_i \gamma} - T_{j,i}^d \right\|_\infty$$

with $T_{j,i}^d$ specified as $T_{j,i}^d = F_{j,i}[G_i]_a$, by designing a detuned H_2-optimal controller for a step reference input using the internal model control (IMC) approach.[4] This is achieved by setting

$$Q_{j,i} = [G_i]_m^{-1} F_{j,i} \tag{29}$$

where $F_{j,i}$ is a lowpass filter of the form

$$F_{j,i} = \left(\frac{\lambda_{j,i}}{s + \lambda_{j,i}} \right)^{n+1}$$

n is the relative degree of the model G_i and $\lambda_{j,i}$ is selected small enough so that $K_{j,i}$ robustly stabilizes G_i.

In the ideal situation where $G_i = G$, the nominal and actual closed-loop transfer functions of the system are equal and are given by the transfer function $F_{j,i}[G_i]_a$. Therefore $\lambda_{j,i}$ is both the nominal and actual closed-loop system bandwidth with a $-3(n+1)$ dB attenuation. In general, $G_i \neq G$ and $\lambda_{j,i}$ serves only as an approximate bandwidth of the actual closed-loop system.

With the controller designed using the above procedure, we shall now show that the transfer function to be identified, $R_{j,i}$, is the product of a known stable proper transfer function and an unknown stable strictly proper transfer function. An analysis of the form of the unknown factor in $R_{j,i}$ indicates how it can be sensibly approximated by a low-order transfer function. We shall first rewrite equation (27) as

$$R_{i,j} = \frac{G - G_i}{1 + Q_{j,i}(G - G_i)} \tag{30}$$

Then we can obtain, after substituting equations (28) and (29) into equation (30) and performing some algebraic manipulations,

$$R_{j,i} = \{[G_i]_m(s+\lambda_{j,i})^n\}\left\{\frac{(s+\lambda_{j,i})(d_{G_i}n_G - d_G n_{G_i})}{d_{K_{j,i}}d_G + n_{K_{j,i}}n_G}\right\} \tag{31}$$

Note that equation (31) can also be written as

$$R_{j,i} = \tilde{R}_{j,i}\hat{R}_{j,i} \tag{32}$$

where

$$\tilde{R}_{j,i} = [G_i]_m(s+\lambda_{j,i})^n \tag{33}$$

is a *known* stable proper transfer function and, other than the factor $s+\lambda_{j,i}$ in the numerator,

$$\hat{R}_{j,i} = \frac{(s+\lambda_{j,i})(d_{G_i}n_G - d_G n_{G_i})}{d_{K_{j,i}}d_G + n_{K_{j,i}}n_G} \tag{34}$$

is an *unknown* stable strictly proper transfer function that depends on the unknown transfer function G. Therefore the problem of identifying $R_{j,i}$ has become one of identifying its unknown factor $\hat{R}_{j,i}$. We shall summarize this important result in the following theorems.

Theorem 3

Consider a plant which has an unknown stable strictly proper transfer function G and a model with a known stable strictly proper transfer function G_i. If G and G_i have no zeros along the imaginary axis of the s-plane and

$$G_i = [G_i]_m[G_i]_a$$

where $[G_i]_m$ is the minimum phase factor of G_i and $[G_i]_a$ is the allpass factor of G_i, then with

$$Q_{j,i} = [G_i]_m^{-1}F_{j,i}$$

and

$$F_{j,i} = \left(\frac{\lambda_{j,i}}{s+\lambda_{j,i}}\right)^{n+1}$$

where n is the relative degree of G_i, the controller

$$K_{j,i} = \frac{Q_{j,i}}{1 - Q_{j,i}G_i}$$

will robustly stabilize G_i for all sufficiently small values of $\lambda_{j,i} \geqslant 0$.

Proof. See Chaps 4 and 5 of Reference 4. □

Theorem 4

Let the controller be designed according to the conditions stated in Theorem 3; then the unknown stable strictly proper transfer function to be identified,

$$R_{j,i} = \frac{G - G_i}{1 + Q_{j,i}(G - G_i)}$$

can be factorized as

$$R_{j,i} = \tilde{R}_{j,i}\hat{R}_{j,i}$$

where $\hat{R}_{j,i}$ is an unknown stable strictly proper transfer function to be identified and $\tilde{R}_{j,i}$ is a known stable proper transfer function given by

$$\tilde{R}_{i,j} = [G_i]_m(s + \lambda_{j,i})^n$$

where $\lambda_{j,i}$ is the nominal closed-loop system bandwidth with a $-3(n+1)$ dB attenuation.

Furthermore, the order and the relative degree (rel deg) of the transfer function $\hat{R}_{j,i}$ are respectively given by

$$\{\text{order of } \hat{R}_{j,i}\} = \{\text{order of } G\} + \{\text{order of } G_i\} - (M+N) + 1$$

$$\text{rel deg}\{\hat{R}_{j,i}\} = \min(\text{rel deg}\{G\}, \text{rel deg}\{G_i\})$$

where M is the number of common zeros in G and G_i and N is the number of common poles in G and G_i.

Proof. See Appendix II. □

Remarks

(i) Note that the factorization of $R_{j,i}$ given in Theorem 4 is naturally induced by the IMC[4] controller design procedure that we have adopted.

(ii) The poles of $\hat{R}_{j,i}$ are the poles of $T_{j,i}$, the actual closed-loop transfer function of the system.

(iii) It is important to note that $\hat{R}_{j,i} = 0$ if and only if $G = G_i$.

(iv) The order of $\hat{R}_{j,i}$ depends on the order of G, which is an unknown.

(v) Although Hansen's approach enables us to obtain an unbiased estimate of the transfer function $\hat{R}_{j,i}$, it should be noted that $\hat{R}_{j,i}$ has more parameters to be estimated than G. Furthermore, since the order, hence the number of parameters to be identified in $\hat{R}_{j,i}$, increases while the magnitude of $\hat{R}_{j,i}$ decreases with the stages of iteration, we would expect that under noisy conditions the system identification problem will become harder as the iteration process progresses. There is an obvious analogy in the windsurfing situation. The better is the skill of a windsurfer, the harder it will be for him/her to improve his/her skill further. In fact, it will take a long time under extreme conditions to improve his/her skill. In the system identification problem for $\hat{R}_{j,i}$ the interpretation is that strong probing signals and a long record of measurements are necessary to achieve even a slight improvement if the closed-loop already has good performance and large bandwidth.

Since we do not know the order of $\hat{R}_{j,i}$ *a priori* and since we are going to identify $\hat{R}_{j,i}$ (actually $R_{j,i}$) and update G_i to G_{i+1} when the step response of the actual closed-loop system exhibits unacceptable oscillations and/or overshoots (associated with model uncertainties), we expect $\hat{R}_{j,i}$ to have complex conjugate poles. Therefore the transfer function which serves as an approximation of $\hat{R}_{j,i}$ has to have an order of at least two. Moreover, since the smallest possible relative degree of a strictly proper transfer function is one and the relative degree of G is unknown, we have to assume that the relative degree of $\hat{R}_{j,i}$ could be one. It was shown in equation (21) that the system identification problem is to find

$$r_{j,i} = \arg\min_{\sigma} \| X_{j,i}Y_{j,i}(R_{j,i} - \sigma)\|_\infty \tag{35}$$

If we define

$$r_{j,i} = \tilde{R}_{j,i}\hat{r}_{j,i} \tag{36}$$

where $\hat{r}_{j,i}$ is an unknown second-order stable strictly proper transfer function, then by substituting equations (32) and (36) into equation (35), we can show that the system identification problem becomes one of finding

$$\hat{r}_{j,i} = \arg \min_{\phi} \| X_{j,i}Y_{j,i}\tilde{R}_{j,i}(\hat{R}_{j,i} - \phi)\|_\infty \tag{37}$$

Therefore, for the purpose of identifying $\hat{R}_{j,i}$, the signal model can be obtained by appropriately modifying equation (22) and is given by

$$\beta_1 = \hat{R}_{j,i}\alpha_2 + v \tag{38}$$

where

$$\alpha_2 = \tilde{R}_{j,i}\alpha_1 \tag{39}$$

and α_1, β_1 and v have been defined previously. Notice that the signals β_1 and α_2 in the model described by equation (38) can easily be generated, using known filters, from the control input u, the measured output y and the reference input r_1.

Remarks

(i) Since $Y_{j,i}$ is the nominal sensitivity function of the closed-loop system, we immediately see that the frequency shaping in the identification criterion given by equation (37) will force the updated model to have small modelling error at the edge of the closed-loop bandwidth where the nominal sensitivity function cannot be made small by the controller $K_{j,i}$.

(ii) It is important to ensure that the input is sufficiently exciting when we are carrying out a system identification experiment.

(iii) Under noisy conditions the signals to be used in the system identification process should be appropriately lowpass filtered. In a discrete time implementation this can be accomplished by an anti-aliasing filter.

(iv) When updating the model using the equation

$$G_{i+1} = G_i + \frac{r_{j,i}}{1 - r_{j,i}Q_{j,i}}$$

the order of the model may increase. To prevent the model order from increasing indefinitely, we use a frequency-weighted balanced truncation scheme[19] to reduce the order of G_{i+1}. Specifically, we find

$$\hat{G}_{i+1} = \arg \min_{\eta} \left\| \frac{G_{i+1}K_{j,i}}{1 + G_{i+1}K_{j,i}} - \frac{\eta K_{j,i}}{1 + \eta K_{j,i}} \right\|_\infty$$

where $\hat{G}_{i+1}$ is the reduced-order model. If the model order is restricted to m, the controller will be at most of order $m + 1$ (see controller design equations given in Theorem 3). In this way the controller complexity will be limited.

7. SIMULATION RESULTS

With reference to Figure 3, we shall present some simulation results of applying the windsurfer approach to the control of a system with

$$G(s) = \frac{9}{(s+1)(s^2+0{\cdot}06s+9)}, \qquad H(s) = 1$$

and e as zero-mean disturbance with a constant energy density of 0·0025 from 0 to 100 Hz. We first summarize the procedure in the following algorithm.

Step 1

Set $G_i = G_0$, where G_0 is the transfer function of an initial model of the plant.

Step 2

Factorize G_i as

$$G_i = [G_i]_m [G_i]_a$$

where $[G_i]_m$ is the minimum phase factor of G_i with a relative degree of n and $[G_i]_a$ is the associated allpass factor of G_i.

Step 3

For $j = 1$ find

$$K_{j,i} = \frac{Q_{j,i}}{1 + Q_{j,i} G_i}$$

with

$$Q_{j,i} = [G_i]_m^{-1} F_{j,i}$$

where the parameter $\lambda_{j,i}$ in the transfer function

$$F_{j,i} = \left(\frac{\lambda_{j,i}}{s + \lambda_{j,i}}\right)^{n+1}$$

is chosen such that $K_{j,i}$ robustly stabilizes G_i in the sense that the filtered (noisy) step response of the actual closed-loop system has, at most, few oscillations and/or overshoots. Stop here if such a robust stabilizing controller cannot be found. Also stop here if the robust stabilizing controller results in a closed-loop system which meets the specified bandwidth. Otherwise proceed to the next step.

Step 4

Let $j = j + 1$ and set $\lambda_{j,i} = \lambda_{j-1,i} + \varepsilon$ for small $\varepsilon > 0$ and redesign the controller $K_{j,i}$ using the equations given in Step 3. Stop here if the design produces a robust stabilizing controller with the closed-loop system satisfying the specified bandwidth. Otherwise repeat this step if $K_{j,i}$ robustly stabilizes G_i; else proceed to the next step.

Step 5

Perform control-relevant system identification to obtain $\hat{f}_{j,i}$. For this purpose we apply an algorithm such as least squares to obtain an estimate $f_{j,i}$ of $\hat{R}_{j,i}$ which satisfies

$$\beta_1 = \hat{R}_{j,i}\alpha_2 + v$$

This depends on using the signals

$$\beta_1 = Y_{j,i}(y - G_i u), \qquad \alpha_2 = \tilde{R}_{j,i}Y_{j,i}X_{j,i}r_1$$

(We actually used discrete time samples of β_1 and α_2 and an output error algorithm to construct a strictly causal second-order estimate from which a continuous time strictly proper $\hat{f}_{j,i}$ was obtained.) Using $\hat{f}_{j,i}$, the model is updated via the following set of equations:

$$\tilde{R}_{j,i} = [G_i]_m(s + \lambda_{j,i})^n, \qquad r_{j,i} = \tilde{R}_{j,i}\hat{f}_{j,i}, \qquad G_{i+1} = G_i + \frac{r_{j,i}}{1 - r_{j,i}Q_{j,i}}$$

Step 6

If G_{i+1} is stable, find the reduced-order model

$$\hat{G}_{i+1} = \arg\min_{\eta} \left\| \frac{G_{i+1}K_{j,i}}{1 + G_{i+1}K_{j,i}} - \frac{\eta K_{j,i}}{1 + \eta K_{j,i}} \right\|_\infty$$

Otherwise stop here.

Step 7

Set $G_i = \hat{G}_{i+1}$ and return to Step 2.

Remarks

(i) In the algorithm, system identification has to be carried out when

$$\| T_{N,i} - \bar{T}_{N,i} \|_\infty$$

is no longer small. Broadly speaking, this will correspond to a significant difference between the designed nominal performance (depending on G_i and $K_{N,i}$) and the actual performance (depending on G and $K_{N,i}$). In particular, the observed step response may exhibit many more oscillations and/or overshoots than the designed values. This is not of course the same thing as guaranteeing that the H_∞ error above has became large, but neither is it unrelated. To be more precise, we define the peak gain of a system whose transfer function is T by

$$\| T \|_1 = \sup_{\| w \|_\infty \neq 0} \frac{\| Tw \|_\infty}{\| w \|_\infty}$$

This is also equal to the *total variation* of the system's unit-step response[23] (roughly the sum of all consecutive peak-to-valley differences in the unit-step response). It can be shown[24] that if T is a stable strictly proper transfer function, then

$$\| T \|_\infty \leqslant \| T \|_1 \leqslant 2p \| T \|_\infty$$

where p is the order of the transfer function T. Now we consider the peak error

$$\| T_{N,i} - \bar{T}_{N,i} \|_1$$

Since

$$\| T_{N,i} - \bar{T}_{N,i} \|_1 \geqslant \| T_{N,i} \|_1 - \| \bar{T}_{N,i} \|_1$$

then if the observed step response of $T_{N,i}$ exhibits many more oscillations and/or overshoots than the designed step response of $\bar{T}_{N,i}$, we would expect

$$\| T_{N,i} \|_1 \gg \| \bar{T}_{N,i} \|_1$$

and hence

$$\| T_{N,i} - \bar{T}_{N,i} \|_1 \gg \varepsilon, \quad \varepsilon > 0$$

Since the peak gain also provides a *loose* lower bound for the H_∞ gain, it is likely that

$$\| T_{N,i} - \bar{T}_{N,i} \|_\infty$$

becomes large when the observed actual step response exhibits many more oscillations and/or overshoots than the desired one. This explains why, in the simulation, the models are updated whenever the filtered (noisy) actual step response exhibits *unacceptable* oscillations and/or overshoots.

(ii) The algorithm used to obtain an estimate $\hat{r}_{j,i}$ of $\hat{R}_{j,i}$ cannot be expected to give an optimal H_∞ estimate. However, note that efficient algorithms for performing H_∞ system identification are still lacking and the corresponding theory is still not well understood.[25–27]

(iii) Since the stability robustness of the closed-loop system for each $\lambda_{j,i}$ has to be checked by using step response testing, the method is not an on-line procedure. In fact, at this stage of development it is an off-line iterative identification and control design procedure.

The simulation results are presented in Figures 6–8. We start with an initial model which has the transfer function

$$G_0 = \frac{0{\cdot}8}{s + 1{\cdot}2}$$

In all these figures the graphs on the left show the noisy unit-step responses of the actual closed-loop systems and those on the right show the corresponding lowpass-filtered signals. Graphs (a) and (b) of Figure 6 show the responses of the actual closed-loop system with a nominal bandwidth of $0{\cdot}1$ rad s^{-1}. Note that overshoots and oscillations are absent for the response in graph (b). Graphs (c) and (d) of Figure 6 are for a nominal closed-loop bandwidth of $0{\cdot}5$ rad s^{-1}. Note that the response in graph (d) is oscillatory and any attempt to increase the nominal closed-loop bandwidth further is likely to lead to instability. At this stage it is necessary to improve the accuracy of the model if we wish to increase the nominal closed-loop bandwidth further. To ensure that the signals are sufficiently exciting, low-amplitude sinusoids in the relevant frequency range are superimposed on the unit-step input just prior to system identification. The responses are shown in graphs (a) and (b) of Figure 7. The updated model has a transfer function

$$G_1 = \frac{0{\cdot}062528s^2 - 0{\cdot}33968s + 10{\cdot}279}{s^3 + 1{\cdot}2801s^2 + 9{\cdot}1173s + 10{\cdot}324}$$

The updated model G_1 is used to redesign a nominal closed-loop system with a bandwidth of $0{\cdot}51$ rad s^{-1} and the responses are shown in graphs (c) and (d) of Figure 7. By comparing graph (d) of Figure 7 with that of Figure 6, we observe that the response no longer has oscillations. We also notice that the rise time in graph (d) of Figure 7 is about twice that in

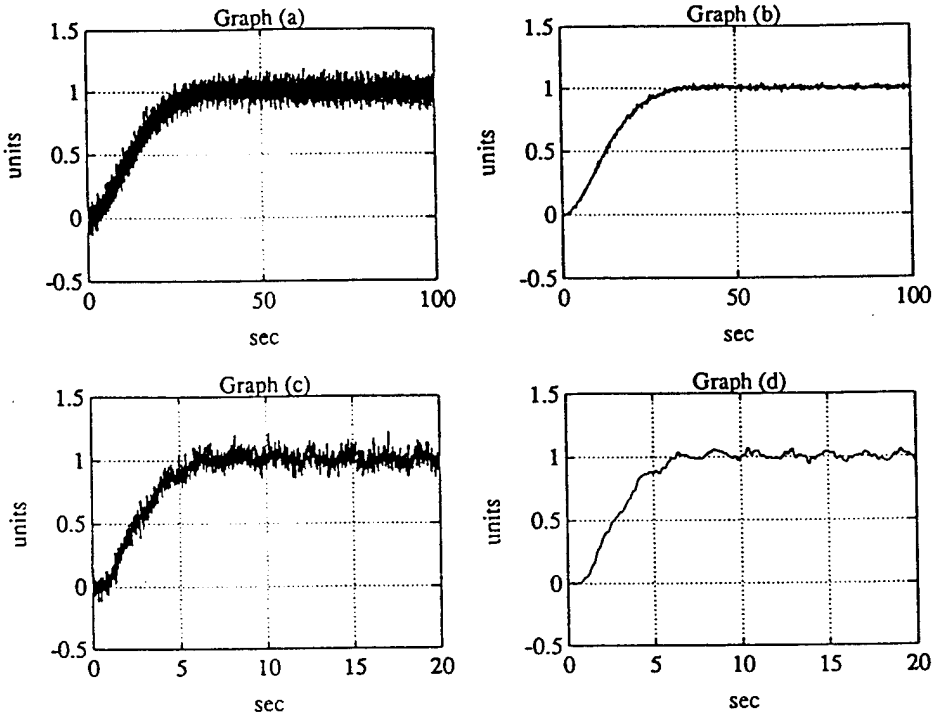


Figure 6. Simulation results 1

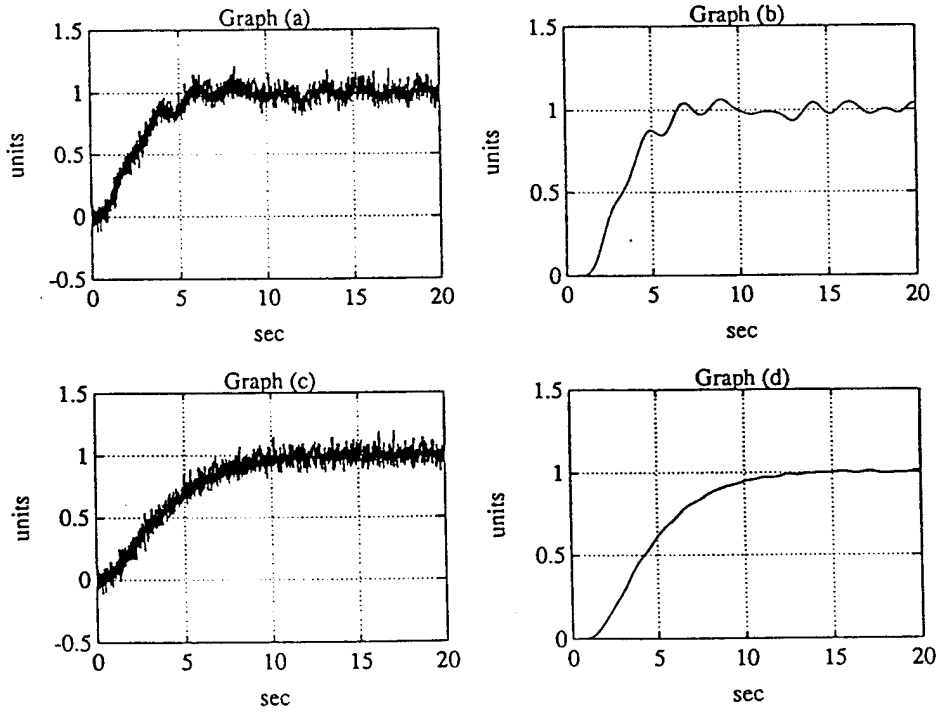


Figure 7. Simulation results 2

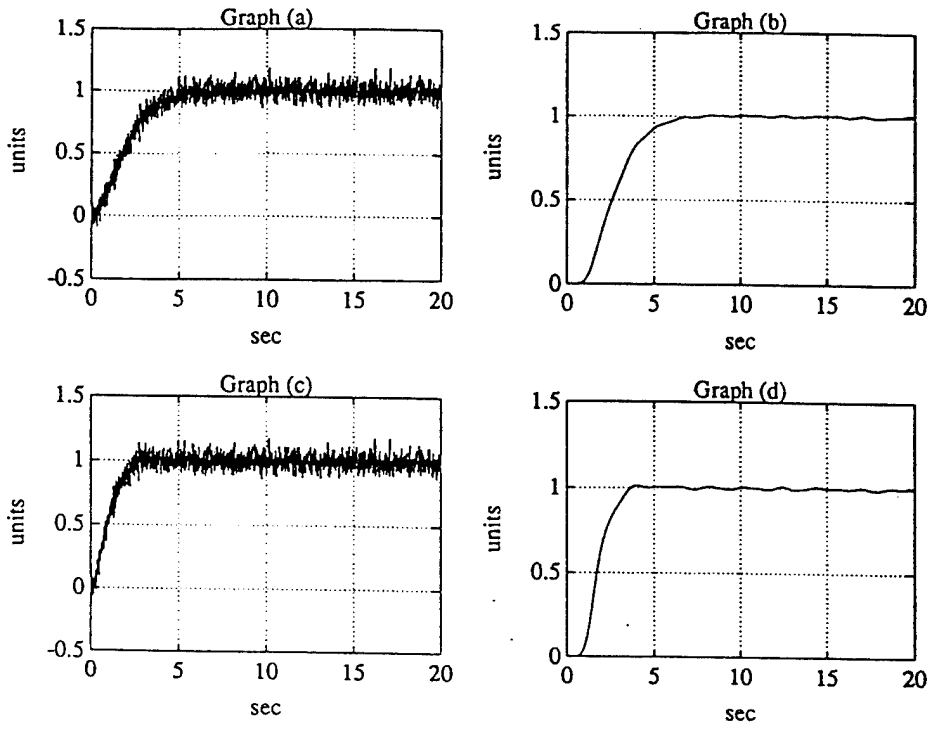


Figure 8. Simulation results 3

graph (d) of Figure 6. Since both G_0 and G_1 have the same relative degree $n = 1$, we would expect graph (d) of Figure 6 and graph (d) of Figure 7 to be similar to the unit-step response of the nominal closed-loop transfer function $[0{\cdot}5/(s+0{\cdot}5)]^2$. By comparing with the computed unit-step response of the transfer function $[0{\cdot}5/(s+0{\cdot}5)]^2$, we have verified that graph (d) of Figure 7 is very close to the desired one. If we continue to increase the nominal closed-loop bandwidth of the system, we obtain the responses shown in Figure 8, where graphs (a) and (b) are for a bandwidth of 1 rad s^{-1} and graphs (c) and (d) are for a bandwidth of 2 rad s^{-1}.

The frequency responses of G, G_0 and G_1 are presented in Figure 9. Notice that, compared with G_0, the updated model G_1 has effectively captured the effects of the poorly damped resonance of the plant.

For the purpose of comparison we present in Figures 10 and 11 the corresponding results obtained using the procedure described in Reference 10. Recall that, as we have mentioned in Section 1, these are obtained under noiseless conditions using rational function approximations (in the H_∞ sense) of the plant instead of identified models. It is also important to emphasize that, instead of strictly proper controllers, proper but non-strictly proper controllers are used in the procedure described in Reference 10.

To facilitate comparison, we adopt the same initial model

$$G_0 = \frac{0{\cdot}8}{s+1{\cdot}2}$$

Graphs (a) and (b) of Figure 10 show the unit-step responses of the actual closed-loop system for nominal closed-loop bandwidths of 0·02 and 0·04 rad s^{-1} respectively. Note that graph (b)

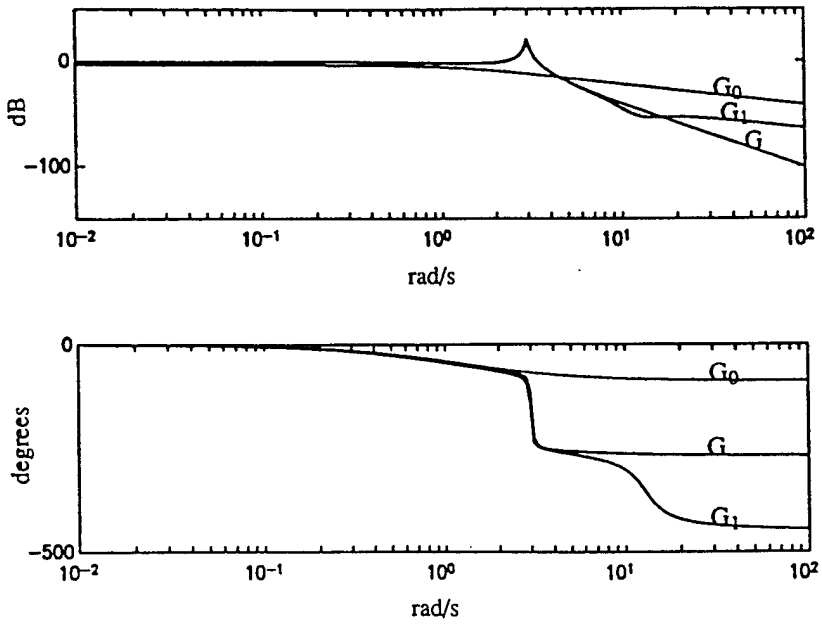


Figure 9. Frequency responses of models and plant

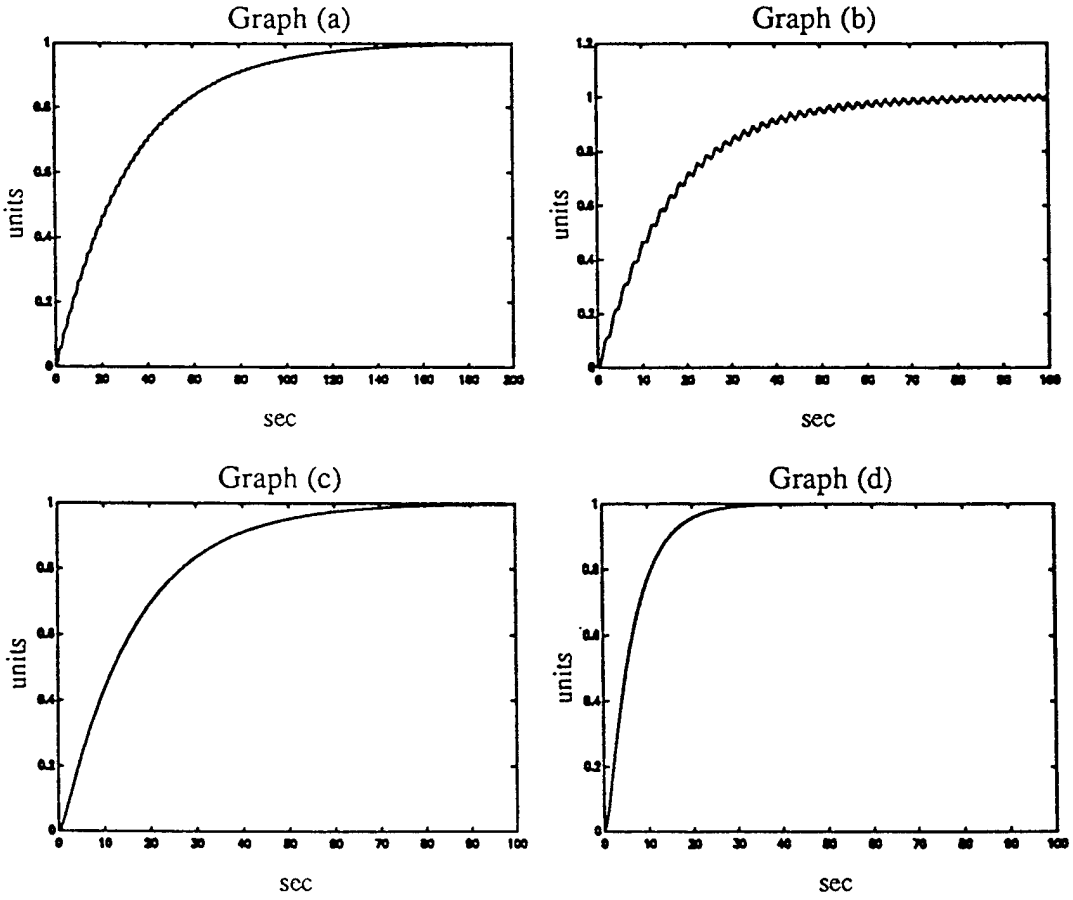


Figure 10. Simulation results using rational approximations of the plant

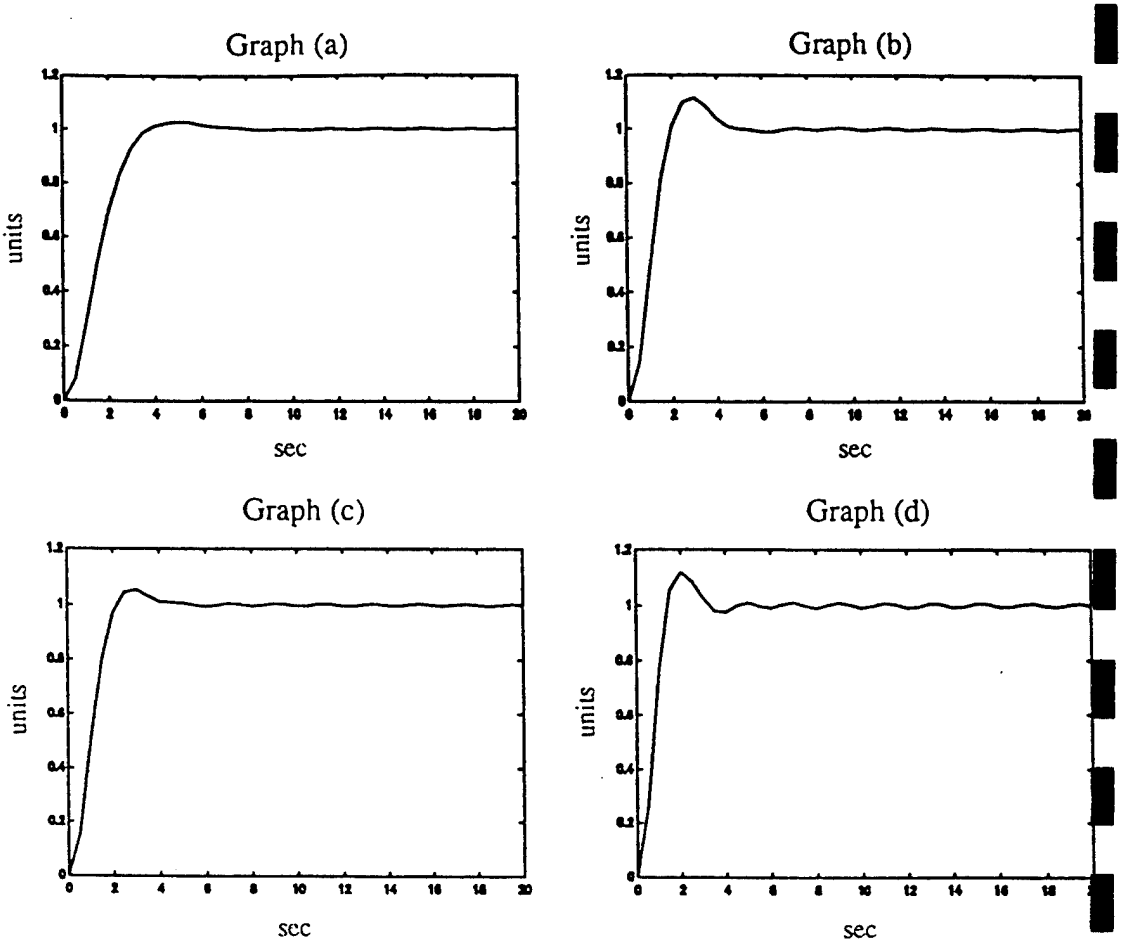


Figure 11. Simulation results using rational approximations of the plant

of Figure 10 shows significant oscillations. After the model is updated to

$$G_1 = \frac{-0{\cdot}4094s^2 + 2{\cdot}0572s + 7{\cdot}175}{s^3 + 1{\cdot}3027s^2 + 8{\cdot}9908s + 10{\cdot}6411}$$

using the procedure described in Reference 10, the unit-step response of the actual closed-loop system for a nominal closed-loop bandwidth of $0{\cdot}04\ \text{rad}\,\text{s}^{-1}$ is improved and it is shown in graph (c) of Figure 10. Graph (d) of Figure 10 shows the unit-step response when the nominal closed-loop bandwidth is increased to $0{\cdot}1\ \text{rad}\,\text{s}^{-1}$. Graphs (a) and (b) of Figure 11 are obtained for nominal closed-loop bandwidths of $0{\cdot}5$ and $1\ \text{rad}\,\text{s}^{-1}$ respectively when the model is G_1. If the model is improved to

$$G_2 = \frac{-0{\cdot}40612s^2 + 0{\cdot}80196s + 6{\cdot}3884}{s^2 + 1{\cdot}0977s^2 + 8{\cdot}882s + 9{\cdot}3027}$$

the unit-step response is given by graph (c) of Figure 11 for a nominal closed-loop bandwidth of $1\ \text{rad}\,\text{s}^{-1}$. When the nominal closed-loop bandwidth is increased to $2\ \text{rad}\,\text{s}^{-1}$, the unit-step response is as shown in graph (d) of Figure 11.

All else being equal, we would expect the noiseless situations to give better results than the noisy conditions. However, by comparing the results given in Figures 6–8 with those given in Figures 10 and 11, we observed that, overall, the results given in Figures 6–8 appear to be better than those given in Figures 10 and 11. Therefore we can conclude that strictly proper controllers are less sensitive to high-frequency model uncertainties and hence require less frequent model updates when we attempt to increase the nominal closed-loop bandwidth of the system. This is important, because, as we have mentioned before, under noisy conditions the system identification process is becoming progressively difficult and it is advantageous to be able to have infrequent but accurate model updates.

8. DISCUSSION AND CONCLUSIONS

We have reviewed in Section 1 the strength and weakness of both the traditional adaptive control and the robust control design methods. These methods should be able to complement each other and there should be natural ways in which they could be blended harmoniously. We proposed that one of the possible ways is by the windsurfer approach which was first mentioned in Reference 13. We have shown, by simulation, that by starting with a (crude) initial model of the plant and a (small-bandwidth) robustly stabilizing controller, the bandwidth of the closed-loop system can be increased progressively through an iterative control-relevant system identification and control design procedure. We shall highlight the following points which we believe are reasons for the success of the approach.

(i) The use of control-relevant frequency weighting in the system identification criterion.
(ii) Updating of the model when the effects of its error are no longer small in the closed-loop response. This will ensure that model uncertainties are emphasized in the correct range of frequencies.
(iii) The controller designed by using the IMC method always has integral action. Therefore it is insensitive to model uncertainties at low frequencies provided that the gain of the model at low frequencies is of the right sign.
(iv) The use of strictly proper controllers to reduce the required number of model updates through identifying the $\hat{R}_{j,i}$ transfer function, a difficult task under noisy conditions.
(iv) The controller designed by using the IMC method induces a natural factorization in the parametrization of the unknown transfer function of the plant. This enables the system identification problem to be solved efficiently.

It is natural in any discussion of adaptive or iterative design to raise the question of convergence. We consider that the work of this paper implicitly established the following practical convergence results, which are confirmed by all the simulations.

1. Assume that a certain closed-loop bandwidth can be achieved for a known stable plant. In the situation where the same plant is imperfectly known, it is possible to adapt the closed-loop system by the method described such that the same closed-loop bandwidth is achieved, assuming that the noise is not so great as to preclude satisfactory identification.
2. For a specified (fixed) closed-loop bandwidth which is achievable with the real plant, the actual closed-loop behaviour approaches very closely the nominal closed-loop behaviour.

In conclusion, we would like to emphasize that only the case of a stable plant with stable models is considered in this preliminary investigation. We hope to address the following problems in the near future.

(i) How to come up with an initial stabilizing controller. Simple cases where the plants have single or double poles at the origin have already been addressed.
(ii) The extension of the method to deal with unstable plants or models.
(iii) Use of orthogonalized exponential in the system identification procedure such that it becomes a convex optimization problem.
(iv) To prove that the algorithm actually converges in some sense.
(v) To study other control design methods in the context of the windsurfer philosophy.

APPENDIX I: PROOF OF THEOREM 2

Since the controller

$$K_{j,i} = \frac{X_{j,i}}{Y_{j,i}} \tag{40}$$

stabilizes the model

$$G_i = \frac{N_i}{D_i}, \qquad N_i X_{j,i} + D_i Y_{j,i} = 1 \tag{41}$$

then solving equations (40) and (41) simultaneously, we get

$$X_{j,i} = \frac{K_{j,i}}{D_i + N_i K_{j,i}} \tag{42}$$

$$Y_{j,i} = \frac{1}{D_i + N_i K_{j,i}} \quad \text{or} \quad Y_{j,i} = \frac{1 - N_i X_{j,i}}{D_i} \tag{43}$$

Substituting $X_{j,i}$ and $Y_{j,i}$ into

$$G_{i+1} = \frac{N_i + r_{j,i} Y_{j,i}}{D_i - r_{j,i} X_{j,i}}$$

will result in

$$G_{i+1} = G_i + \frac{r_{j,i}}{D_i (D_i - r_{j,i} X_{j,i})}$$

Solving for $r_{j,i}$, we get

$$r_{j,i} = \frac{D_i^2 (G_{i+1} - G_i)}{1 + D_i X_{j,i} (G_{i+1} - G_i)} \tag{44}$$

From Figure 3 in Section 5, with $r_2 = 0$, we can write for the closed-loop system

$$y = \frac{G K_{j,i}}{1 + G K_{j,i}} r_1 + \frac{1}{1 + G K_{j,i}} He, \qquad u = \frac{K_{j,i}}{1 + G K_{j,i}} r_1 - \frac{K_{j,i}}{1 + G K_{j,i}} He$$

Therefore we can write the equation

$$\beta = D_i y - N_i u$$

as

$$\beta = \frac{D_i K_{j,i} (G - G_i)}{1 + G K_{j,i}} r_1 + \frac{D_i (1 + G_i K_{j,i})}{1 + G K_{j,i}} He \tag{45}$$

and the equation

$$\alpha = X_{j,i} r_1$$

as

$$\alpha = \frac{K_{j,i}}{D_i (1 + G_i K_{j,i})} r_1 \tag{46}$$

If we form the output error defined by

$$\varepsilon - \beta - r_{j,i}\alpha \tag{47}$$

then by substituting equations (44)–(46) into equation (47) and using the expression for $X_{j,i}$ given by equation (42), we can obtain

$$\varepsilon = \frac{D_i(1+G_iK_{j,i})K_{j,i}(G-G_{i+1})}{(1+GK_{j,i})(1+G_{i+1}K_{j,i})} r_1 + \frac{D_i(1+G_iK_{j,i})}{1+GK_{j,i}} He \tag{48}$$

Since equation (43) can also be written as

$$Y_{j,i} = \frac{1}{D_i(1+G_iK_{j,i})}$$

it is clear that if we define the filtered output error as

$$\xi = Y_{j,i}\varepsilon$$

then

$$\xi = \frac{K_{j,i}(G-G_{i+1})}{(1+GK_{j,i})(1+G_{i+1}K_{j,i})} r_1 + \frac{1}{1+GK_{j,i}} He$$

or

$$\xi = \left(\frac{GK_{j,i}}{1+GK_{j,i}} - \frac{G_{i+1}K_{j,i}}{1+G_{i+1}K_{j,i}}\right) r_1 + \frac{1}{1+GK_{j,i}} He \qquad \square$$

APPENDIX II: PROOF OF THEOREM 4

Using the notations established in Section 6, we have

$$R_{j,i} = \frac{G-G_i}{1+Q_{j,i}(G-G_i)}, \qquad Q_{j,i} = \frac{K_{j,i}}{1+G_iK_{j,i}}$$

Therefore we can write

$$R_{j,i} = \frac{1+G_iK_{j,i}}{1+GK_{j,i}} (G-G_i) \tag{49}$$

We also have

$$G_i = [G_i]_m [G_i]_a$$

where

$$[G_i]_m = \frac{\tilde{n}_{G_i}\Pi_i (z_i^* + s)}{d_{G_i}}, \qquad [G_i]_a = \frac{\Pi_i (z_i - s)}{\Pi_i (z_i^* + s)}$$

Since $Q_{j,i} = [G_i]_m^{-1} F_{j,i}$, we can rewrite the equation

$$K_{j,i} = \frac{Q_{j,i}}{1 - Q_{j,i}G_i}$$

as

$$K_{j,i} = \frac{d_{G_i}n_F}{\tilde{n}_{G_i}[\Pi_i (z_i^* + s)d_{F_{j,i}} - \Pi_i (z_i - s)n_{F_{j,i}}]}$$

Hence we can write

$$1 + G_iK_{j,i} = \frac{\Pi_i (z_i^* + s)d_{F_{j,i}}}{\Pi_i (z_i^* + s)d_{F_{j,i}} - \Pi_i (z_i - s)n_{F_{j,i}}} \tag{50}$$

By substituting equation (50) into equation (49) and noting that

$$d_{F_{j,i}} = (s + \lambda_{j,i})^{n+1}, \qquad 1 + GK_{j,i} = \frac{d_G d_{K_{j,i}} + n_G n_{K_{j,i}}}{d_G d_{K_{j,i}}}, \qquad G - G_i = \frac{d_{G_i}n_G - d_G n_{G_i}}{d_G d_{G_i}}$$

we obtain

$$R_{j,i} = \tilde{R}_{j,i}\hat{R}_{j,i}$$

where

$$\tilde{R}_{j,i} = [G_i]_m(s + \lambda_{j,i})^n$$

is a *known* stable proper transfer function and, other than the factor $s + \lambda_{j,i}$ in the numerator,

$$\hat{R}_{j,i} = \frac{(s + \lambda_{j,i})(d_{G_i}n_G - d_G n_{G_i})}{d_{K_{j,i}}d_G + n_{K_{j,i}}n_G}. \tag{51}$$

is an *unknown* stable strictly proper transfer function.

To obtain the results on the order and relative degree of $\hat{R}_{j,i}$ we shall write

$$[G_i]_m = \frac{\eta_i(s)}{\pi_i(s)}, \qquad [G_i]_a = \frac{\rho_i(-s)}{\rho_i(s)}$$

where each of the polynomials $\eta_i(s)$, $\pi_i(s)$ and $\rho_i(s)$ has degree r, $n + r$ and m respectively. We can then obtain

$$K_{j,i} = \frac{\lambda_{j,i}^{n+1}\pi_i(s)\rho_i(s)}{\eta_i(s)[(s + \lambda_{j,i})^{n+1}\rho_i(s) - \lambda_{j,i}^{n+1}\rho_i(-s)]}$$

If we also write G as

$$G = \frac{\alpha(s)}{\beta(s)}$$

where $\alpha(s)$ has degree p and $\beta(s)$ has degree q, then by substituting all these into equation (51), we get

$$\hat{R}_{j,i} = \frac{(s + \lambda_{j,i})[\alpha(s)\pi_i(s)\rho_i(s) - \eta_i(s)\rho_i(-s)\beta(s)]}{\beta(s)\eta_i(s)[(s + \lambda_{j,i})^{n+1}\rho_i(s) - \lambda_{j,i}^{n+1}\rho_i(-s)] + \lambda_{j,i}^{n+1}\alpha(s)\pi_i(s)\rho_i(s)} \tag{52}$$

By counting the degrees of the resulting numerator and denominator polynomials of $\hat{R}_{j,i}$ given by equation (52), the required results are established immediately. □

REFERENCES

1. Horowitz, I., *Synthesis of Feedback Systems*, Academic, New York, 1963.
2. Goodwin, G. C., and K. S. Sin, *Adaptive Filtering, Prediction and Control*, Prentice-Hall, Englewood Cliffs, NJ, 1984.
3. Doyle, J. C., *Lecture Notes — ONR/Honeywell Workshop on Advances on Multivariable Control*, Minneapolis, MN, 1984.
4. Morari, M., and E. Zafiriou, *Robust Process Control*, Prentic-Hall, Englewood CLiffs, NJ, 1989.
5. Rohrs, C. E., L. Valavani, M. Athans and G. Stein, 'Robustness of continuous-time adaptive control algorithms in the presence of unmodeled dynamics', *IEEE Trans. Automatic Control*, AC-30, 881–889 (1985).
6. Anderson, B. D. O., R. R. Bitmead, C. R. Johnson, Jr., P. V. Kokotovic, R. L. Kosut, I. M. Y. Mareels, L. Praly and B. D. Riedle, *Stability of Adaptive Systems: Passivity and Averaging Analysis*, MIT Press, Cambridge, MA, 1986.
7. Åström, K. J., and B. Wittenmark, *Adaptive Control*, Addison-Wesley, Reading, MA, 1989.
8. Kosut, R. L., M. K. Lau and S. P. Boyd, 'System identification for robust control design', *Proc. ECC '91*, Grenoble, July 1991, pp. 1384–1389.
9. Zang, Z., and R. R. Bitmead, 'Transient bounds for adaptive control systems', *Proc. CDC '90*, Honolulu, HW, December 1990, pp. 2724–2729.
10. Lee, W. S., B. D. O. Anderson, R. L. Kosut and I. M. Y. Mareels, 'On adaptive robust control and control-relevant system identification', *Proc. ACC '92*, Chicago, IL, June 1992.
11. Zang, Z., R. R. Bitmead and M. Gevers, 'H_2 iterative model refinement and control robustness enhancement', *Proc. CDC '91*, Brighton, December 1991, pp. 279–284.
12. Zang, Z., R. R. Bitmead and M. Gevers, 'Disturbance rejection: on-line refinement of controllers by closed loop modelling', *Proc. ACC '92*, Chicago, IL, June 1992.
13. Anderson, B. D. O., and R. L. Kosut, 'Adaptive robust control: on-line learning', *Proc. CDC '91*, Brighton, December 1991, pp. 297–298.
14. Hansen, F. R., and G. F. Franklin, 'On a fractional representation approach to closed-loop experiment design', *Proc. ACC '88*, Atlanta, GA, June 1988, pp. 1319–1320.

15. Hansen, F. R., G. F. Frankin and R. L. Kosut, 'Closed-loop identification via the fractional representation: experiment design', *Proc. ACC '89*, Pittsburgh, PA, June 1989, pp. 1422–14276.
16. Hansen, F. R., 'A fractional representation approach to closed-loop system identification and experiment design', *Ph.D. Dissertation*, Stanford University, March 1989.
17. Schrama, R. J. P., 'Accurate identification for control: the necessity of an iterative scheme', *IEEE Trans. Automatic Control*, **AC-37**, 991–994 (1992).
18. Freudenberg, J. S., and D. P. Looze, 'Right half plane poles and zeros and design tradeoffs in feedback systems', *IEEE Trans. Automatic Control*, **AC-30**, 555–565 (1985).
19. Anderson, B. D. O., and Y. Liu, 'Controller reduction: concepts and approaches', *IEEE Trans. Automatic Control*, **AC-34**, 802–812 (1989).
20. Vidyasagar, M. *Control System Synthesis: A Factorization Approach*, MIT Press, Cambridge, MA, 1985.
21. Ljung, L., *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, NJ, 1987.
22. Kokotovic, P. V., H. K. Khalil and J. O'Reilly, *Singular Perturbation Methods in Control: Analysis and Design*, Academic, London, 1986.
23. Boyd, S. P., and C. H. Barratt, *Linear Controller Design*, Prentice-Hall, Englewood Cliffs, NJ, 1991.
24. Boyd, S. P., and J. C. Doyle, 'Comparison of peak and RMS gains for discrete-time systems', *Syst. Control Lett.*, **9**, 1–6 (1987).
25. Helmicki, A. J., C. A. Jacobson and C. N. Nett, 'Control oriented system identification: a worst-case/deterministic approach in H_∞' *IEEE Trans. Automatic Control*, **AC-36**, 1163–1176 (1991).
26. Parker, P. J., and R. R. Bitmead, 'Adaptive frequency response identification', *Proc. CDC '87*, Los Angeles, CA, December 1987, pp. 348–353.
27. Partington, J. R., 'Robust indentification and interpolation in H_∞', *Int. J. Control*, **54**, 1281–1290 (1991).

A.3 Unbiased least-squares estimates with structure incompatibilities

H. Aling and R.L. Kosut, *Proc. 1993 ECC*, Groningen, The Netherlands, 29 June-2 July, 1993.

Unbiased least squares estimates in the presence of structure incompatibilities

Henk Aling *and Robert L. Kosut †‡

Abstract : When the system identification problem is posed in a stochastic framework, model errors are defined as the sum of a bias and a variance term. The bias term is caused by incompatibilities between the model structure used with the identification and that of the true system. In particular, the use of ARX models almost invariably leads to bias since associated noise model has no zeros. Results are known on the unbiasedness of the first n Markov parameters of n-th order least squares ARX model estimates for systems that are excited by a white noise input. We will extend these results to the colored input case.

Keywords: Least Squares, Markov Parameter, System Identification, Model Reduction, Unbiased Estimation.

Introduction

Let us assume that we want to identify a model of the following linear, time-invariant discrete time system:

$$y_t = G(z)u_t + H(z)e_t$$

Here, y_t and u_t are the observed output and input sequences, and G and H are transfer functions in the shift operator z :

$$\begin{aligned} G(z) &= G_0 + G_1 z^{-1} + G_2 z^{-1} + \cdots \\ H(z) &= I + H_1 z^{-1} + H_2 z^{-2} + \cdots \end{aligned}$$

$G(z)$ and $H(z)$ are general in the sense that we do not assume finite-dimensionality, or compatibility with an ARX model structure. We will use an n-th order ARX model to identify the system:

$$y_t + A_1 y_{t-1} + \ldots + A_n y_{t-n} = B_0 u_t + \ldots + B_n u_{t-n} + \epsilon_t \quad (1)$$

This model structure is represented in terms of transfer functions as

$$y_t = A(z)^{-1} B(z) u_t + A(z)^{-1} e_t$$

where

$$A(z) = I + \sum_{i=1}^{n} A_i z^{-i} \quad , \quad B(z) = \sum_{i=0}^{n} B_i z^{-i}$$

Since the all-pole noise model structure $A(z)^{-1}$ is generally incompatible with the noise system $H(z)$, any estimate $\hat{G}(z) = \hat{A}(z)^{-1}\hat{B}(z)$ of $G(z)$ will be biased. This bias can be reduced significantly by increasing n , thereby allowing more degrees of freedom to fit the noise model. Thanks to the robustness properties of the least squares method this can oftentimes be done without severe errors due to overfitting. However, if the excitation is poor, large errors in the frequency response of the model may still result in the poorly excited regions. A better approach towards unbiased estimation is given by the following known result, based on the impulse response of the estimated ARX model [4, 2, 5, 6]:

Theorem 1 *Let u_t and e_t both be independently, identically distributed sequences, and let u_t and e_t be uncorrelated. Then, as the number of data points goes to infinity, the first $n+1$ Markov parameters $\hat{G}_0, \ldots, \hat{G}_n$, obtained from the n-th order ARX model, are unbiased estimate of the true Markov parameters $G_0, \ldots, G_n$.*

The surprising property expressed by this theorem is that the first $n+1$ Markov parameters are estimated biasfree *even in the case where the ARX model structure is incompatible with that of the true system.* This has been proved for the noisefree case [4, 2] as well as for the noisy case [5, 6]. We will extend this theorem to the case where u_t is colored. The theorem is particularly useful for reduction of high order ARX models based on the model impulse response [3, 7] .

Definitions and normal equations

We will derive a direct relationship between the Markov parameters of the true system and the ARX model parameters, based on a straightforward evaluation of the normal equations. We define a parameter

*Research Scientist, Integrated Systems Inc., 3260 Jay Street, Santa Clara, CA 95054, USA. Phone: (408) 980 1500, Fax: (408) 980 0400, Email: aling@isi.com.

†Manager Basic Research, Integrated Systems Inc., 3260 Jay Street, Santa Clara, CA 95054, USA.

‡This research was supported by the National Science Foundation under contract no. ISI-9161408

matrix Θ and regression vector ϕ_t by

$$\Theta = \begin{pmatrix} \mathcal{A} & \mathcal{B} \end{pmatrix} = \begin{pmatrix} A_1 & \dots & A_n & B_0 & \dots & B_n \end{pmatrix}$$

$$\phi_t = \begin{pmatrix} -y_{t-1}^T & \dots & -y_{t-n}^T & u_t^T & \dots & u_{t-n}^T \end{pmatrix}^T$$

Hence, we can write (1) as

$$y_t = \Theta\phi_t + \epsilon_t$$

The least square solution is determined by the normal equations

$$\hat{\Theta} X X^T = Y X^T \qquad (2)$$

where Y and X are matrices containing the output and regression data:

$$Y = \begin{pmatrix} y_1 & \dots & y_N \end{pmatrix}, \quad X = \begin{pmatrix} \phi_1 & \dots & \phi_N \end{pmatrix}$$

To derive the relationship between the ARX model parameters and the true Markov parameters (G_i, H_i) we have to introduce some more notation:

$$\begin{aligned} g &= \begin{pmatrix} G_0 & G_1 & G_2 & \dots \end{pmatrix} \\ h &= \begin{pmatrix} I & H_1 & H_2 & H_3 & \dots \end{pmatrix} \\ U_t &= \begin{pmatrix} u_t^T & u_{t-1}^T & u_{t-2}^T & \dots \end{pmatrix}^T \\ E_t &= \begin{pmatrix} e_t^T & e_{t-1}^T & e_{t-2}^T & \dots \end{pmatrix}^T \end{aligned}$$

The input/output relationship can now be written as

$$y_t = gU_t + hE_t \qquad (3)$$

To put this in matrix form for all $t = 1, \dots, N$ we have to add the definitions

$$\mathcal{U} = \begin{pmatrix} U_1 & U_2 & U_3 & \dots & U_N \end{pmatrix}$$

$$\mathcal{E} = \begin{pmatrix} E_1 & E_2 & E_3 & \dots & E_N \end{pmatrix}$$

We can now write (3) for all t as

$$Y = \begin{pmatrix} g & h \end{pmatrix} \begin{pmatrix} \mathcal{U} \\ \mathcal{E} \end{pmatrix}$$

For the regression matrix X we can define

$$\mathcal{G} = \begin{pmatrix} 0 & G_0 & G_1 & \dots & G_{n-1} & \dots \\ 0 & 0 & G_0 & \dots & G_{n-2} & \dots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \dots \\ 0 & 0 & \dots & 0 & G_0 & \dots \end{pmatrix}$$

$$\mathcal{H} = \begin{pmatrix} 0 & I & H_1 & \dots & H_{n-1} & \dots \\ 0 & 0 & I & \dots & H_{n-2} & \dots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \dots \\ 0 & 0 & \dots & 0 & I & \dots \end{pmatrix}$$

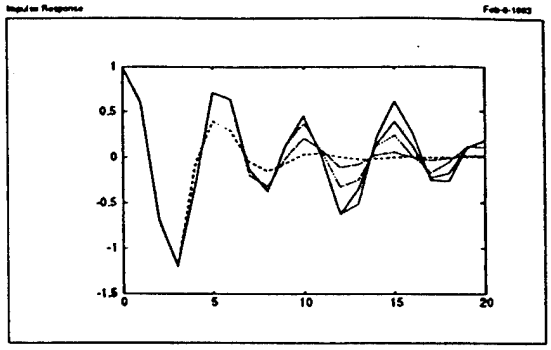


Figure 1: Impulse responses of ARX models, of order 3,6,9,12: white input case.

and express the matrix of regression vectors as

$$X = \begin{pmatrix} -\mathcal{G} & -\mathcal{H} \\ \mathcal{I} & 0 \end{pmatrix} \begin{pmatrix} \mathcal{U} \\ \mathcal{E} \end{pmatrix}$$

Here, $\mathcal{I}$ is an identity matrix extended with zeros to the right. The normal equations are now written compactly as

$$\begin{pmatrix} I & \hat{\mathcal{A}} & -\hat{\mathcal{B}} \end{pmatrix} \begin{pmatrix} g & h \\ \mathcal{G} & \mathcal{H} \\ \mathcal{I} & 0 \end{pmatrix} \cdot \begin{pmatrix} \mathcal{U}\mathcal{U}^T & \mathcal{U}\mathcal{E}^T \\ \mathcal{E}\mathcal{U}^T & \mathcal{E}\mathcal{E}^T \end{pmatrix} \begin{pmatrix} \mathcal{G}^T & \mathcal{I}^T \\ \mathcal{H}^T & 0 \end{pmatrix} = 0 \qquad (4)$$

Unbiased Markov parameter estimation

The first main result is theorem 1 which been stated already. The proof is a straight consequence of the normal equations (4) and is given in the appendix.

As an example, we have estimated models of orders 3, 6, 9 and 12 on simulated data (generated by a single input/single output 4-th order system with white noise as input, and a signal to noise ratio of 1). Their impulse responses are shown in Figure 1; the solid line is the impulse response of the true model. The result is a good illustration of the theory: the impulse responses of all models are identical up to $t = 3$, after which the third order response starts to differ from the rest - we see two separate lines. After $t = 6$ the sixth order response starts to differ so that we see 3 different lines, and so on. The fit of the first $n+1$ Markov parameters is based on asymptotic results and is therefore exact.

A first extension to the non-white input case is given by the following corollary of theorem 1 . The idea is

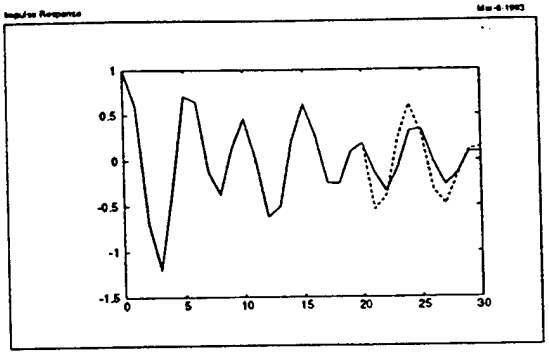


Figure 2: 20-th order ARX model impulse response vs. true model based on prefiltered input.

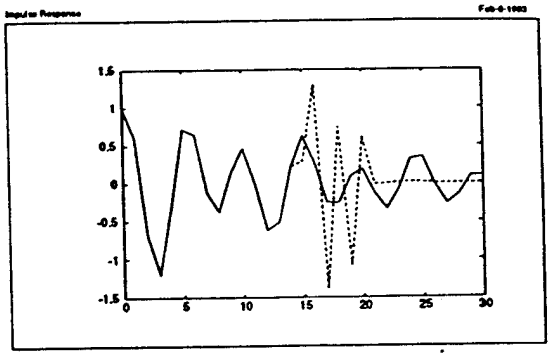


Figure 3: 20-th order ARX model impulse response vs. true model based on standard least squares.

to describe the u_t as the output of a filter with a white sequence w_t as input, and use w_t as input for the ARX model. The proof is given in the appendix.

Corollary 2 *Let $V(z)$ be the spectral density function of u_t and let it have the spectral factorization $V(z) = W(z)\Sigma_{uu}W(z^{-1})^T$ [1] . Define $w_t = W(z)^{-1}u_t$, and $\hat{\Gamma}(z) = \hat{A}(z)^{-1}\hat{B}(z)$ as the n-th order least squares estimate of an ARX model with input w_t and output y_t . Let $\hat{G}(z) = \sum_{i=0}^{\infty}\hat{G}_i z^{-i}$ be defined as $\hat{G}(z) = \hat{\Gamma}(z)W(z)^{-1}$. Then, as the number of data points goes to infinity, $\hat{G}_0, \ldots, \hat{G}_n$ are unbiased estimates of the true Markov parameters $G_0, \ldots, G_n$.*

The corollary is illustrated by figure 2. The figure shows the comparison between the impulse response of a 20-th order ARX model (dashed line) estimated from data generated by a 4-th order model (solid line) observed in additive white noise (output error structure, as opposed to ARX which is an equation error structure). The order r of the input filter with both poles and zeros was chosen to be 6, and the signal to noise ratio was equal to 1.

This result is a straightforward consequence of theorem 1 and requires the availability of the signal w_t . Due to the definition of $\hat{G}(z)$, the plant model order will generally be higher than n , the order of the combined plant/filter system. It would be nice if we could relate the the Markov parameters of a n-th order ARX model with input u_t directly to the true Markov parameters, as in theorem 1 . The following theorem proves that this is possible indeed, if the input is an auto regressive filtered white noise. The proof is given in the appendix.

Theorem 3 *Let u_t be a quasi-stationary sequence with spectral density function $V(z)$ of the form $V(z) = W(z)\Sigma_{uu}W(z^{-1})^T$, where $\Sigma_{uu} = \Sigma_{uu}^T > 0$ and where $W(z)$ is an auto regressive finite-dimensional transfer function described by*

$$W(z) = D(z)^{-1}, \quad D(z) = I + \sum_{i=1}^{r} D_i z^{-i}$$

Let e_t be an independently, identically distributed sequence uncorrelated with u_t . Assume that $r < n$. Then, as the number of data points goes to infinity, the first $n + 1 - r$ Markov parameters $\hat{G}_0, \ldots, \hat{G}_{n-r}$, obtained from the n-th order ARX model, are unbiased estimates of the true Markov parameters $G_0, \ldots, G_{n-r}$.

The theorem is illustrated by figure 3. The figure shows the comparison between the impulse response of a 20-th order ARX model (dashed line) estimated from data generated by a 4-th order model (solid line) observed in additive white noise. The order r of the autoregressive input filter was chosen to be 6, and the signal to noise ratio was equal to 1. Therefore, the impulse responses are identical up to the $n+1-r = 15$ -th Markov parameter.

Conclusions

Despite the incompatibility of the ARX model structure with that of the underlying system, it is possible to use least squares to obtain unbiased model estimates. This can be done in two ways which leads to either $n+1$ unbiased Markov parameters if a prefiltered white input is used, or $n+1-r$ unbiased Markov parameters in case the input is the output of an r-th order Auto Regressive filter driven by white noise. The results derived in this paper are based on an asymptotic analysis.

Then, we may postmultiply (9) by the $(n+1)\times(n+1-r)$ block matrix

$$\begin{pmatrix} I & & \\ D_1^T & \ddots & \\ \vdots & \ddots & I \\ D_r^T & & D_1^T \\ & \ddots & \vdots \\ & & D_r^T \end{pmatrix}$$

This leads to

$$\begin{pmatrix} I & A \end{pmatrix} \begin{pmatrix} \Delta_0 & \dots & \Delta_n \\ & \ddots & \vdots \\ & & \Delta_0 \end{pmatrix} \begin{pmatrix} I & & \\ C_1^T & \ddots & \\ \vdots & \ddots & I \\ C_r^T & & C_1^T \\ & \ddots & \vdots \\ & & C_r^T \end{pmatrix} = 0 \tag{10}$$

In the all-pole case, we have $C_1 = \dots = C_r = 0$. Then, (10) leads to $\Delta_0 = \dots = \Delta_{n-r} = 0$. Since $\hat{G}(z) = \Delta(z)\Sigma_{uu}^{-1}D(z)$, by straightforward multiplication of coefficients we obtain

$$G_i = \hat{G}_i \quad (i = 0, \dots, n-r)$$

which proves the theorem. From (10) it is evident what the problems are when $W(z)$ has zeros as well. Then, it is not straightforward how a similar result can be obtained. We leave that case as a direction for future research.

References

[1] B.D.O. Anderson, J.B. Moore : *Optimal filtering*, ISBN 0-13-638122-7, Prentice Hall, New Jersey, 1979.

[2] Y. Inouye, *Approximation of multivariable linear systems with impulse response and autocorrelation sequences*, Automatica, vol. 19 no. 3, pp. 265-277, 1983.

[3] S.Y. Kung: *A new identification and model reduction algorithm with singular value decomposition*, Proc. 12th Asilomar Conf. Circuits and Systems, Pacific Grove, California, nov. 1978, pp.705-714.

[4] C.T. Mullis and R.A. Roberts, *The use of second order information in the approximation of linear systems*, IEEE-ASSP, vol.24 no.3, june 1976.

[5] H. Swaanenburg, W. Schinkel, G.A. van Zee, O.H. Bosgra: *Practical aspects of industrial multivariable process identification*, Preprints 7th IFAC/IFORS Symposium, York, pp. 201-206, 1985.

[6] P.M.J. Van den Hof, P.H.M. Janssen: *Some asymptotic properties of multivariable models identified by equation error techniques*, IEEE-AC, vol. 32, pp. 89-92, 1987.

[7] H.P. Zeiger and A.J. McEwen, *Approximate linear realization of given dimension via Ho's algorithm*, IEEE-AC, vol. 19, p. 153, 1974.

Appendix

Proof of theorem 1

According to our assumptions, we may set $\mathcal{U}\mathcal{E}^T$ and $\mathcal{E}\mathcal{U}^T$ to zero in expectation. The right block column of (4) then leads to

$$E\{\begin{pmatrix} I & \hat{A} & -\hat{B} \end{pmatrix} \begin{pmatrix} \mathcal{G}\mathcal{U}\mathcal{U}^T\mathcal{I}^T \\ \mathcal{G}\mathcal{U}\mathcal{U}^T\mathcal{I}^T \\ \mathcal{I}\mathcal{U}\mathcal{U}^T\mathcal{I}^T \end{pmatrix}\} = 0 \qquad (5)$$

By definition of Markov parameters we also have

$$\begin{pmatrix} I & \hat{A} & -\hat{B} \end{pmatrix} \begin{pmatrix} \hat{\mathcal{G}}\mathcal{U}\mathcal{U}^T\mathcal{I}^T \\ \hat{\mathcal{G}}\mathcal{U}\mathcal{U}^T\mathcal{I}^T \\ \mathcal{I}\mathcal{U}\mathcal{U}^T\mathcal{I}^T \end{pmatrix} = 0 \qquad (6)$$

Let us define the (i,j)-th block element of $\mathcal{U}\mathcal{U}^T$ as

$$V_{j-i} = E\{\sum_t u_{t-i}u_{t-j}^T\}$$

Note that this term depends on $j-i$ only, due to the stationarity assumption. Subtracting (5) from (6), taking expectations and introducing $\tilde{G}_k$ for the difference $\hat{G}_k - G_k$, we get

$$\begin{pmatrix} I & \hat{A} \end{pmatrix} \left(\begin{array}{ccc|ccc} \tilde{G}_0 & \dots & \tilde{G}_n & \tilde{G}_{n+1} & \dots \\ & \ddots & \vdots & \vdots & \\ & & \tilde{G}_0 & \tilde{G}_1 & \dots \end{array}\right) \cdot$$

$$\cdot \left(\begin{array}{ccc} V_0 & \dots & V_n \\ \vdots & \ddots & \vdots \\ V_{-n} & \dots & V_0 \\ \hline V_{-n-1} & \dots & V_{-1} \\ \vdots & & \vdots \end{array}\right) = 0 \qquad (7)$$

If u_t is white noise, this implies that $V_k = N\Sigma_{uu}\delta_k$. Substituting that, it is easily seen that (7) leads to $\tilde{G}_k = 0$ $(k = 0,\dots,n)$ which proves the theorem.

Proof of corollary 2

Let the joint plant/filter transfer function be defined as $\Gamma(z) = G(z)W(z)$, and let $W(z)$ have no zeros on the complex unit circle. Thus, $W(z)$ has a stable inverse and $G(z) = \Gamma(z)W(z)^{-1}$. Subtracting $\hat{G}(z) = \hat{\Gamma}(z)W(z)^{-1}$ from this, we get $(\hat{G}(z) - G(z))W(z) = \hat{\Gamma}(z) - \Gamma(z)$. Since by theorem 1 $\hat{\Gamma}_i = \Gamma(i)$ $(i = 0,\dots,n)$, the coefficients of $z^0,\dots,z^{-n}$ of the Taylor series expansion of this term are zero. Thus, $\hat{G}_i = G_i$ for $i = 0,\dots,n$.

Proof of theorem 3

In case the input is colored, the expectation of the term $\mathcal{U}\mathcal{U}^T$ will not be a block diagonal matrix but a block Toeplitz matrix. Assume that the input spectral density function is given by

$$V(z) = W(z)\Sigma_{uu}W(z)^* \ , \quad W(z) = I + \sum_{i=1}^{r} W_i z^{-i}$$

where $(W(z), \Sigma_{uu})$ represent the standard innovations representation of $V(z)$ according to the spectral factorization theorem [1]. Then, (7) becomes

$$\begin{pmatrix} I & \hat{A} \end{pmatrix} \left(\begin{array}{ccc|ccc} \tilde{G}_0 & \dots & \tilde{G}_n & \tilde{G}_{n+1} & \dots \\ & \ddots & \vdots & \vdots & \\ & & \tilde{G}_0 & \tilde{G}_1 & \dots \end{array}\right) \cdot$$

$$\cdot \left(\begin{array}{ccc|ccc} I & \dots & W_n & W_{n+1} & \dots \\ & \ddots & \vdots & \vdots & \\ & & I & W_1 & \dots \\ \hline & & & I & \dots \\ & & & & \ddots \end{array}\right) \mathrm{diag}(\Sigma_{uu}).$$

$$\cdot \left(\begin{array}{ccc} I & & \\ \vdots & \ddots & \\ W_n^T & \dots & I \\ \hline W_{n+1}^T & \dots & W_1^T \\ \vdots & & \vdots \end{array}\right) = 0 \qquad (8)$$

or, defining $\Delta(z) = \tilde{G}(z)W(z)\Sigma_{uu} = \sum_{i=0}^{\infty}\Delta_i z^{-i}$:

$$\begin{pmatrix} I & \hat{A} \end{pmatrix} \left(\begin{array}{ccc|ccc} \Delta_0 & \dots & \Delta_n & \Delta_{n+1} & \dots \\ & \ddots & \vdots & \vdots & \\ & & \Delta_0 & \Delta_1 & \dots \end{array}\right) \cdot$$

$$\cdot \left(\begin{array}{ccc} I & & \\ \vdots & \ddots & \\ W_n^T & \dots & I \\ \hline W_{n+1}^T & \dots & W_1^T \\ \vdots & & \vdots \end{array}\right) = 0 \qquad (9)$$

Now suppose $W(z) = D(z)^{-1}C(z)$ with

$$D(z) = I + \sum_{i=1}^{r} D_i z^{-i} \ , \quad C(z) = I + \sum_{i=1}^{r} C_i z^{-i}$$

A.4 On some key issues in the windsurfer approach to adaptive control

W.S. Lee, B.D.O. Anderson, I.M.Y. Mareels, and R.L. Kosut, *Automatica*, Vol. 31, No. 11, pp.1619-1636, 1995. Appeared first in *Proc. SYSID 94*, 10th IFAC Symposium on System Identification, Copenhagen, Denmark, 4-6 July, 1994.

Pergamon

0005-1098(95)00092-5

Automatica, Vol. 31, No. 11, pp. 1619–1636, 1995
Elsevier Science Ltd
Printed in Great Britain
0005-1098/95 $9.50 + 0.00

On Some Key Issues in the Windsurfer Approach to Adaptive Robust Control

W. S. LEE,† B. D. O. ANDERSON,‡ I. M. Y. MAREELS§ and R. L. KOSUT¶

In the context of iterative identification and control, where the control objective is to maximize the bandwidth of the closed-loop system while achieving step tracking, we provide a methodology that achieves this goal for stable systems.

Key Words—Adaptive control; robust control; internal model control; identification.

Abstract—We examine a number of crucial questions that arise in the windsurfer approach to adaptive robust control. Considerations are limited to the case where the plant is stable and has no zeros on the imaginary axis. The key conclusion is that, given a strictly proper stable model of a strictly proper stable plant, we can improve the performance robustness of the closed-loop system through the windsurfer approach if the plant and the existing model have no unstable zeros within the designed closed-loop bandwidth and if the deterioration in performance robustness caused by increasing the closed-loop bandwidth results in a sufficiently high signal-to-noise ratio for a certain closed-loop output error. Situations that may cause the iterative identification and control design process to terminate prematurely are identified. A simulation example is used to illustrate the results discussed.

1. INTRODUCTION

1.1. *Background and objectives*

A new adaptive control paradigm known as the *windsurfer approach* was first introduced in Anderson and Kosut (1991). The objective of this approach is to increase the bandwidth of a closed-loop system, if possible to a specified value through an *iterative identification and control design procedure*, given that the initial model of the plant may involve significant error in the high-frequency region. Furthermore, as the closed-loop bandwidth is being increased, the closed-loop frequency response is to be kept approximately flat in the passband so that the closed-loop transient response is not too oscillatory or having excessive peak overshoot.

Iterative identification and control design is a topic of growing interest (see e.g. Anderson and Kosut, 1991; Zang *et al.*, 1991; Schrama, 1992; Schrama and Van den Hof, 1992; Lee *et al.*, 1993; Partanen and Bitmead, 1993). Although these schemes are different in detail and have been proposed for achieving different control objectives, they have all originated from the awareness that, in any model-based control design task, the model serves no other purpose than that of designing a controller. It is therefore not surprising that the identification criterion adopted in each of these schemes is determined by the respective control performance criterion. Furthermore, the controllers employed in each of these schemes are designed on the basis of models (except possibly the first one) identified from data obtained under closed-loop conditions. An historical perspective on as well as a tutorial introduction to the *joint design of identification and control* may be found in Gevers (1993).

A scheme for the windsurfer approach was presented in Lee *et al.* (1993). It was demonstrated by simulations that the bandwidth of a closed-loop system can be increased by the iterative applications of the internal model control (IMC) method (see Morari and Zafiriou, 1989) and a closed-loop system identification procedure pioneered by Hansen (1989).

* Received 6 May 1994; received in final form 7 June 1995. This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Petros Ioannou under the direction of Editor C. C. Hang. Corresponding author Dr Iven M. Y. Mareels. Tel. +61 6 2493378; Fax +61 6 2490506; E-mail Iven.Mareels@anu.edu.au.

† Department of Electrical and Electronic Engineering, Victoria University of Technology, PO Box 14428 MMC, Victoria 3000, Australia.

‡ Department of Systems Engineering and Cooperative Research Centre for Robust and Adaptive Systems, Australian National University, Canberra, ACT 0200, Australia.

§ Department of Engineering Faculty of Engineering and Information Technology Australian National University Canberra, ACT 0200, Australia.

¶ Integrated Systems, Inc., 3260 Jay St., Santa Clara, CA 95054, U.S.A.

In this paper we examine a number of crucial questions that arise in this approach. When can one redesign the controller and expand the closed-loop bandwidth, without re-identifying; When should one re-identify? What does one want to identify in the re-identification process? What can one identify in the re-identification process? How can an identified model be verified against the desired purpose? Will re-identification always lead to improved closed-loop performance? Attention is restricted to strictly proper stable plants with no finite zeros on the imaginary axis. Extensions to more general situations are currently under investigation (see Campi *et al.*, 1994).

1.2. *Structure of the paper*

In Section 2 we describe the IMC design state of the windsurfer approach and show that it is safe to increase the designed closed-loop bandwidth gradually if the plant is stabilized by the existing controller. Section 2 also introduces some of the key concepts and notation used in the paper. Properties of good models for the windsurfer approach are established in Section 3. The control-relevant system identification method employed by the windsurfer approach is described in Section 4. Conditions necessary for identifying a good model and methods for verifying experimentally that an identified model is suitable for the desired purpose (or otherwise) will be given. In Section 5 we study mechanisms that may influence the iterative identification and control design of the windsurfer approach. Situations that may lead to the premature termination of the iterative process will be indicated. In Section 6 two methods for model validation are described. A procedure for the identification of a better model while avoiding the potential danger of causing instability in the actual closed-loop system will then be suggested. Conditions under which the performance robustness of a closed-loop system can be improved through the windsurfer approach are discussed in Section 7. A simulation example is presented in Section 8. We conclude the paper in Section 9.

2. PRELIMINARIES

In this section we describe a closed-loop system where, on the basis of a strictly proper stable model, a sequence of controllers is designed for a strictly proper stable plant. This description also introduces some of the key concepts and notation used in this paper. In particular, Section 2.1 outlines the IMC method in the manner that it is applied in the control design step of the windsurfer approach. The concepts of nominal performance and robust stability are then introduced. In Section 2.2 it is shown that we can increase the designed closed-loop bandwidth of the system, while maintaining the stability of the actual closed-loop system, if the increment is sufficiently small. We conclude this section with a definition of performance robustness relevant to the windsurfer approach.

2.1. *Controller design in the windsurfer approach*

The IMC method is applied in the control design step of the windsurfer approach where the reference input is a step function. Although the IMC method is generally applicable to the case where the plant and the models are not necessarily stable, we restrict ourselves to the case where the plant and the models are strictly proper and stable. In this situation the designed closed-loop bandwidth of the system is determined by a single design parameter.

Consider a closed-loop system as shown in Fig. 1, where G is the transfer function of a strictly proper stable plant. A sequence of such models (identified from data obtained in closed-loop) eventuates in the windsurfer approach. We use G_i to denote the ith member in the sequence of strictly proper stable models $\{G_0, G_1, G_2, \ldots\}$. On the basis of G_i, a finite sequence of controllers $\{K_i^0, K_i^1, \ldots, K_i^{f_i}\}$ is designed such that, while keeping the closed-loop frequency responses approximately flat within the pass bands, the corresponding closed-loop bandwidths form an increasing sequence $\{\lambda_1^0, \lambda_i^1, \ldots, \lambda_0^{f_i}\}$. Note that we shall in general use K_i to denote one of the controllers in the sequence $\{K_i^0, K_i^1, \ldots, K_i^{f_i}\}$ when it is immaterial to the discussion which particular controller is involved. Figure 1 shows that

$$K_i = \frac{Q_i}{1 - G_i Q_i}, \tag{1}$$

with Q_i defined in the IMC method by

$$Q_i = [G_i]_m^{-1} F_i, \tag{2}$$

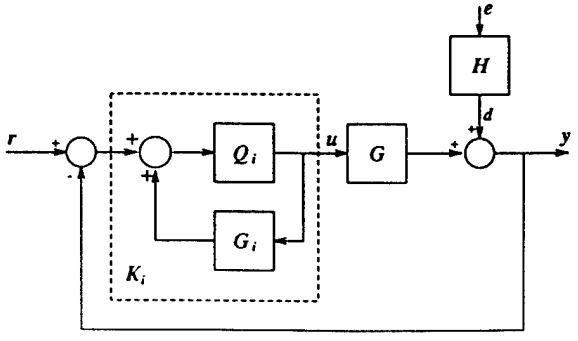


Fig. 1. Internal model control structure.

where $[G_i]_m$ is the *minimum-phase factor* of G_i, and

$$F_i = \left(\frac{\lambda_i}{s+\lambda_i}\right)^n, \quad \lambda_i > 0, \tag{3}$$

is a suitable IMC filter when the model is stable and when the reference input is a step function. The integer n of the IMC filter is chosen such that Q_i is proper. The design parameter λ_i must be chosen such that the *actual closed-loop transfer function*

$$T_i = \frac{TK_i}{1+GK_i} \tag{4}$$

is stable.

It can be shown that the *designed closed-loop transfer function*

$$\bar{T}_i = \frac{G_iK_i}{1+G_iK_i}$$

can be written as

$$\bar{T}_i = G_iQ_i \tag{5}$$

or

$$\bar{T}_i = F_i[G_i]_a, \tag{6}$$

where $[G_i]_a$ is the all-pass factor associated with G_i.

Remarks.

- Since $[G_i]_a(j\omega)$ does not affect the magnitude of $\bar{T}_i(j\omega)$, it is clear that $|\bar{T}_i(j\omega)|$ is flat in its passband and λ_i is the *designed closed-loop bandwidth* with an attenuation of $-3n$ dB.
- Note that the system becomes open-loop when λ_i approaches zero. Since G is stable, it is always possible to make T_i stable by choosing a sufficiently small λ_i.

Although the designed closed-loop transfer function $\bar{T}_i$ is always well behaved, the actual closed-loop transfer function T_i may become unstable when λ_i is too large. We introduce the following definitions.

Definition 1. The designed closed-loop system involving K_i has *robust stability* if the stability of $\bar{T}_i$ implies the stability of T_i. K_i is said to *robustly stabilize* G_i.

Definition 2. For any two closed-loop systems designed by the method described in this subsection, we say that the one with a larger value of λ_i has a better *nominal performance.*

2.2. *Improving nominal performance while maintaining stability*

In Section 2.1 we have indicated that a closed-loop system may not have robust stability when λ_i becomes too large. Since the objective of the windsurfer approach is to increase the closed-loop bandwidth to a specificed value, the following question appears naturally:

- When can the closed-loop bandwidth be increased *with safety*; that is, *without losing robust stability*, while retaining the use of the model G_i?

To answer the above question, we recall from Lee *et al.* (1993) that if T_i (corresponding to λ_i) is stable then there exists a strictly proper transfer function R_i such that

$$G = G_i + \frac{R_i}{1-Q_iR_i}.$$

It can then be shown that

$$\tilde{T}_i = Q_i(1-\bar{T}_i)R_i, \tag{7}$$

where $\tilde{T}_i = T_i - \bar{T}_i$ is the error in the closed-loop transfer function induced by the error in the model G_i when the designed closed-loop bandwidth is λ_i.

Suppose that the designed closed-loop bandwidth is increased to $\lambda_i' > \lambda_i$; then, corresponding to λ_i', we can write

$$\tilde{T}_i' = Q_i'(1-\bar{T}_i')R_i',$$

where

$$R_i' = \frac{R_i}{1+[G_i]_m^{-1}(F_i'-F_i)R_i}.$$

Since Q_i' and $\bar{T}_i'$ are stable by design, $\tilde{T}_i'$ and T_i' are stable if and only if R_i' is stable. However, R_i' is stable if $\lambda_i' - \lambda_i > 0$ is *sufficiently small.* Hence we have the following conclusion.

Conclusion 1. We can increase the designed closed-loop bandwidth *cautiously* if the existing closed-loop system has robust stability.

Remark. Note that even when T_i is stable, its response to the reference input could be significantly different from that of $\bar{T}_i$, if, relative to the frequencies where G_i has significant errors, λ_i is not sufficiently small. To address this issue, we now introduce the concept of robust performance that is relevant to the windsurfer approach.

Definition 3. With respect to the given reference

input r and a specified finite (usually suitably small) $\sigma > 0$, the closed-loop system is said to have *robust performance* with designed closed-loop bandwidth λ_i if and only if

$$J_i \stackrel{\text{def}}{=} \|v_i\|_2^2 \leq \sigma,$$

where $v_i = \tilde{T}_i r$ is the *tracking error.*

Remarks.

- Robust stability is necessary for robust performance, but it is not sufficient.
- It is important to note that a closed-loop system may have high nominal performance (large λ_i) but poor robust performance ($J_i > \sigma$), and vice versa.
- For a model with significant modelling errors in the high-frequency region, the closed-loop system can be designed to have good robust performance if the designed closed-loop bandwidth is sufficiently small.
- While λ_i is being increased, a stage can be reached (before the occurrence of instability) where, because of the modelling errors associated with G_i making a significant contribution to J_i, the performance robustness has deteriorated beyond an acceptable level. At this stage the designed closed-loop bandwidth is $\lambda_i = \lambda_i^f$, and we decide to obtain a more accurate model G_{i+1} before continuing to open up the bandwidth.

3. PROPERTIES OF GOOD MODELS

In Section 2.2 we have concluded that when performance robustness of the closed-loop system has deteriorated beyond an acceptable level, it is necessary to identify a model better than the existing one before the designed closed-loop bandwidth can be increased further.

It is clear from Section 2.2 that we can increase the designed closed-loop bandwidth as long as the closed-loop system has robust performance. Therefore it is natural that, when the closed-loop system looses robust performance, we attept to seek a new model that will allow robust performance of the closed-loop system to be restored through controller redesign (while the designed closed-loop bandwidth remains unchanged). This prompts us to ask the following question.

- What *would we like* to identify, in order that, with the new model, robust performance of the closed-loop system can be improved through controller redesign?

Before we proceed to answer the last question, we should observe that each stage of the windsurfer approach involves an existing model G_i and an updated model G_{i+1}. Since every stage of the iteration proceeds in a similar fashion, it suffices to discuss only the stage where $i = 0$. Therefore we shall denote the existing model by G_0 and the updated model by G_1. This system of notation will carry over to all transfer functions and signals involved in the following discussions.

Suppose that G_1 is identified when λ_0 has reached λ_0^f. A new controller K_1^0 will then be designed on the basis of G_1 such that λ_1^0 has the same value as λ_0^f. Obviously, we should like $J_1^0 = \|\tilde{T}_1^0 r\|_2^2$ to be small. By using (1)–(5), with appropriate adjustments made to the notation, we can write $\tilde{T}_1^0 = T_1^0 - \bar{T}_1^0$ as

$$\tilde{T}_1^0 = \frac{\dfrac{G - G_1}{G_1}\bar{T}_1^0}{1 + \dfrac{G - G_1}{G_1}\bar{T}_1^0}(1 - \bar{T}_1^0). \tag{8}$$

Clearly, it is necessary that $\tilde{T}_1^0$ be stable. Since $G - G_1$ is unknown, we conclude the following.

Conclusion 2. We *would like* to identify G via a G_1 of sufficient accuracy such that the model G_1 satisfies the sufficient condition of robust stability

$$\left\| \frac{G - G_1}{G_1}\bar{T}_1^0 \right\|_\infty < 1.$$

Furthermore, we observe that the magnitude of the designed sensitivity function $1 - \bar{T}_1^0$ in the right hand side of (8) could approach a magnitude significantly greater than one if G_1 has unstable zeros within the passband of $\bar{T}_1^0 = F_1^0[G_1]_a$. In order that $\tilde{T}_1^0$ have a small magnitude, we require in addition to the above robust stability condition the following.

Conclusion 3. We *would like* to identify G via a G_1 of sufficient accuracy such that

$$\left| \frac{G(j\omega) - G_1(j\omega)}{G_1(j\omega)}\bar{T}_1^0(j\omega) \right|$$

is sufficiently small for all frequencies above the lesser of the passband of $\bar{T}_1^0$ and the smallest critical frequency corresponding to the unstable zeros of $\bar{T}_1^0$.

Remarks.

- Observe that the unstable zeros of $\bar{T}_1^0$ are those of G_1, which, in a situation with good identification, will be those of the plant G.
- If G_1 has unstable zeros located within the passband of $\bar{T}_1^0$, it is likely that there is a range of frequency within the passband of $\bar{T}_1^0$ where the magnitude of the designed sensitivity

function $1-\bar{T}_1^0$ is significantly greater than one. This has the following consequences:

(1) there is a range of frequency within the passband of $\bar{T}_1^0$ where the designed system has poor disturbance rejection and the measurement noise is not well attenuated;

(2) since the magnitude of the designed sensitivity function is the inverse of the distance of the open-loop frequency response curve from the critical point of stability at $-1+j0$, the designed system may have poor stability margins and transient response if the magnitude of the designed sensitivity function is excessively large near the edge of the system passband.

For these reasons, we may not want to increase the designed closed-loop bandwidth λ_1 beyond λ_0^f if G_1 is found to have unstable zeros with critical frequencies within the passband of $\bar{T}_0^f$.

4. SYSTEM IDENTIFICATION IN THE WINDSURFER APPROACH

Notwithstanding the fact that we have established in the last section properties of a good model for the windsurfer approach, it is important to ask the following question.

- What *can we* identify by using the system identfication procedure embedded in the windsurfer approach?

In this section we answer this question in three steps. In Section 4.1 we show that consideration of the control objective leads to a (*closed-loop*) *control-relevant system identification problem* that can be transformed into an open-loop system identification problem. This transformation is achieved by employing an identification framework pioneered by Hansen (1989), where, instead of the plant itself, a strictly proper stable transfer function (to be denoted by R_0^f) that parametrizes the plant is identified. In Section 4.2 we shall show that it is possible to identify R_0^f accurately only if the signal-to-noise ratio of a certain closed-loop output error resulting from the existing controller is high. Furthermore, by recognizing the relation between the signal component of the closed-loop output error and deterioration in robust performance, we can restate the conditions necessary for obtaining an accurate estimate of R_0^f in terms of the level of deterioration in robust performance against the effect of noise disturbance. In Section 4.3 we show how to verify indirectly that an estimate of R_0^f is unbiased.

4.1. *Control-relevant system identification*

It was indicated at the end of Section 2.2 that when the designed closed-loop bandwidth has reached a certain value denoted by λ_0^f, the robust performance measure

$$J_0^f = \|v_0^f\|_2^2$$

associated with the closed-loop system designed on the basis of G_0 would become excessively large. It was shown in Section 3 that at this stage, we like to identify a new model G_1 such that

$$\left|\frac{G(j\omega)-G_1(j\omega)}{G_1(j\omega)}\bar{T}_1^0(j\omega)\right|$$

is sufficiently small in an appropriate frequency range. Unfortunately it is not clear how to process input–output measurements to determine G_1 so that this condition is naturally or automatically satisfied. To overcome this difficulty, we use input–output measurements and possibly the reference input of the stable closed-loop system as shown in Fig. 2 to identify G_1 such that

$$\left\|\left(\frac{GK_0^f}{1+GK_0^f}-\frac{G_1K_0^f}{1+G_1K_0^f}\right)r\right\|_2^2$$

is minimized. This closed-loop identification problem can be transformed into an open-loop identification problem by employing Hansen's (1989) framework of identification. We state this result in the following theorem. It is a special case of Theorem 2 in Lee *et al.* (1993) when the plant G and the model G_0 are stable.

Theorem 1. Let $K_0^f=(1-G_0Q_0^f)^{-1}Q_0^f$ stabilize G and G_0, where Q_0^f is a proper stable transfer function, so that G can be parametrized by a strictly proper stable transfer function R_0^f via

$$G = G_0 + \frac{R_0^f}{1-R_0^fQ_0^f}.$$

Let

$$G_1 = G_0 + \frac{\hat{R}_0^f}{1-\hat{R}_0^fQ_0^f} \tag{9}$$

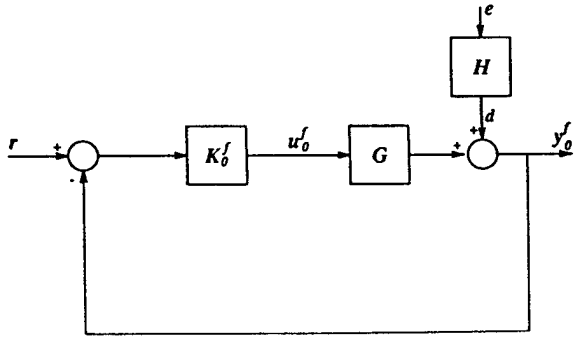


Fig. 2. Closed-loop system just before identification.

be another model stabilized by K_0^f, where $\hat{R}_0^f$ is a strictly proper stable estimate of R_0. Also define

$$\xi_1 = (1 - \bar{T}_0^f)(\beta - \hat{R}_0^f \alpha), \tag{10}$$

where $\alpha = Q_0^f r$, $\beta = y_0^f - G_0 u_0^f$, and u_0^f and y_0^f are respectively the input and output of the plant resulting from the application of K_0^f. Then ξ_1 can be expressed as

$$\xi_1 = \left(\frac{GK_0^f}{1 + GK_0^f} - \frac{G_1 K_0^f}{1 + G_1 K_0^f}\right) r + w_0^f, \tag{11}$$

where

$$w_0^f = (1 - T_0^f) He \tag{12}$$

is the effect of the noise disturbance e on the actual closed-loop output.

Remarks.

- If we define $H = (1 - R_0^f Q_0^f)^{-1} S_0^f$, where S_0^f is a proper stable and inversely stable transfer function, then the actual closed-loop system has Hansen's open-loop representation

$$\beta = R_0^f \alpha + S_0^f e. \tag{13}$$

- From Theorem 1, it is clear that minimizing

$$\left\| \left(\frac{GK_0^f}{1 + GK_0^F} - \frac{G_1 K_0^f}{1 + G_1 K_0^f}\right) r + w_0^f \right\|_2^2$$

with respect to G_1 is equivalent to minimizing

$$\|(1 - G_0 Q_0^f)(\beta - \hat{R}_0^f \alpha)\|_2^2$$

with respect to $\hat{R}_0^f$, provided that G_1 is updated according to

$$G_1 = G_0 + \frac{\hat{R}_0^f}{1 - \hat{R}_0^f Q_0^f}.$$

Note that the appropriate signal model (which has also taken the data filters L into consideration) for this system identification problem is

$$\bar{\beta} = R_0^f \bar{\alpha} + \Psi_0^f e,$$

where

$$\bar{\alpha} = L(1 - \bar{T}_0^f)\alpha,$$

$$\bar{\beta} = L(1 - \bar{T}_0^f)\beta,$$

$$\Psi_0^f = L(1 - \bar{T}_0^f) S_0^f.$$

- Since the 'input' α in (13) (and hence $\bar{\alpha}$) is independent of the noise disturbance e, identifying R_0^f and S_0^f (or equivalently Ψ_0^f) is an *open-loop* identification problem.
- We identify R_0^f and Ψ_0^f using a prediction error method (see Ljung, 1987) as shown in Fig. 3. Data filters L (typically low-pass) are usually employed to shape the bias-distribution of the estimates (which is due to under-modelling) such that the model error is small in the appropriate frequency range.

We can summarise the above discussions as follows.

Conclusion 4. We can transform the closed loop identification of G into an open-loop identification problem for

$$R_0^f = \frac{G - G_0}{1 + Q_0^f (G - G_0)}.$$

4.2. *Accurate identification of* R_0^f

In the following we show that the problem of identifying R_0^f accurately can be solved effectively (using finitely many input–output measurements) if the signal-to-noise ratio of a certain closed-loop output error (to be defined immediately) is high. In particular, the normalized variance for an unbiased estimate of R_0^f is small if the signal-to-noise ratio associated with the closed-loop output error is sufficiently high.

From Fig. 4, we observe that the *closed-loop*

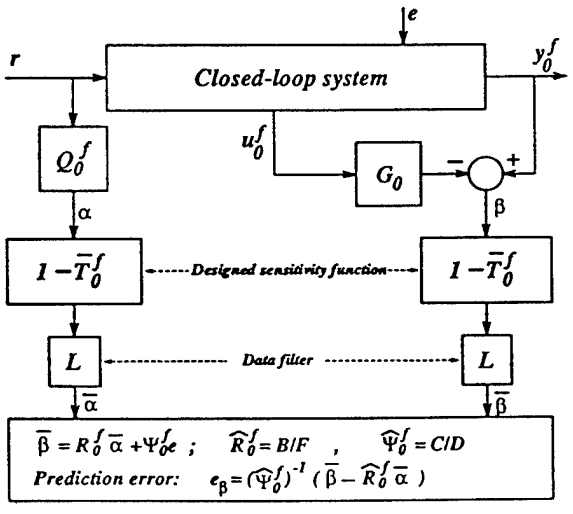


Fig. 3. Identification of R_0^f and Ψ_0^f.

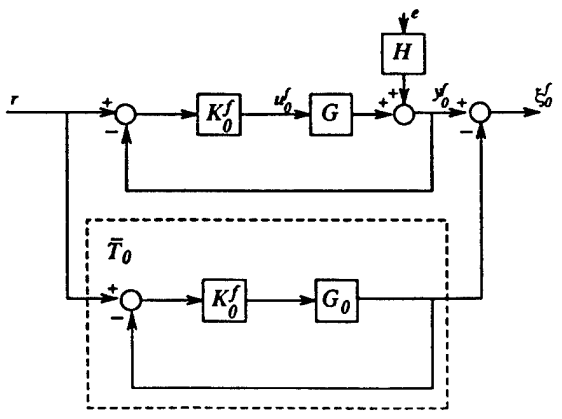


Fig. 4. Closed-loop output error.

output error ξ_0^f is defined as $\xi_0^f = y_0^f - \bar{T}_0^f r$. By substituting the expressions for α and β into (10) and noting that

$$u_0^f = \frac{Q_0^f}{1 - G_0 Q_0^f}(r - y_0^f),$$

we can obtain

$$\hat{R}_0^f = \arg\min_{\rho} \|\xi_0^f - \rho Q_0^f(1 - \bar{T}_0^f)r\|_2^2. \quad (14)$$

Now we can use the fact that $y_0^f = T_0^f r + (1 - T_0^f)He$ to write

$$\xi_0^f = v_0^f + w_0^f, \quad (15)$$

where

$$v_0^f = \bar{T}_0^f r. \quad (16)$$

Remark. Note that the tracking error v_0^f cannot be measured directly. It can only be estimated from the closed-loop output error ξ_0^f.

It is apparent that v_0^f is the signal component in ξ_0^f that carries the useful information about the existing modelling errors under closed-loop condition, and w_0^f is the noise component in ξ_0^f which obstructs the determination of $\hat{R}_0^f$. Therefore we can draw an immediate conclusion.

Conclusion 5. We can identify R_0^f successfully if the signal-to-noise ratio associated with the closed-loop output error resulting from the existing controller K_0^f is high.

We next show that the normalized variance for an unbiased estimate of R_0^f is small in the frequency range where the signal-to-noise ratio associated with the closed-loop output error is sufficiently high.

By substituting (15) and (16) into (14) and noting from (7) that $\bar{T}_0^f = Q_0^f(1 - \bar{T}_0^f)R_0^f$, we can write

$$\hat{R}_0^f = \arg\min_{\rho} \|Q_0^f(1 - \bar{T}_0^f)(R_0^f - \rho)r + w_0^f\|_2^2.$$

In practice, we use sampled input–output data to estimate a discrete-time model for $\hat{R}_0^f$ before converting it to a continuous-time transfer function. We assume that the errors involved in this conversion are negligible. Following Ljung (1987), we write the variance of an unbiased estimate of R_0^f approximately as

$$\mathbf{E}(|\hat{R}_0^f(j\omega) - R_0^f(j\omega)|^2) \sim \frac{m}{M} \frac{\Phi_{w_0^f}(\omega)}{|Q_0^f(j\omega)[1 - \bar{T}_0^f(j\omega)]|^2 \Phi_r(\omega)},$$

where $\Phi_{w_0^f}(\omega)$ is the power spectral density of w_0^f, under the condition that the order of the discrete-time model for $\hat{R}_0^f$ (denoted by m) and the number of data (denoted by M) are large and the ratio m/M is small. Since

$$\Phi_{v_0^f}(\omega) = |Q_0^f(j\omega)[1 - \bar{T}_0^f(j\omega)]R_0^f(j\omega)|^2 \Phi_r(\omega)$$

is the power spectral density of v_0^f, we can write the normalized variance of $\hat{R}_0^f$ as

$$\mathbf{E}\left(\left|\frac{\hat{R}_0^f(j\omega) - R_0^f(j\omega)}{R_0^f(j\omega)}\right|^2\right) \sim \frac{m}{M} \frac{\Phi_{w_0^f}(\omega)}{\Phi_{v_0^f}(\omega)}$$

for the frequencies where $R_0^f(j\omega) \neq 0$.

Remark. For a finite number of data, the normalized variance of $\hat{R}_0^f$ can be small only in the frequency range where the signal-to-noise ratio associated with the closed-loop output error is sufficiently high.

We now summarise the above discussion as follows.

Conclusion 6. We can obtain an unbiased estimate of R_0^f with a small normalized variance in a certain frequency range $\omega_1 \leq \omega \leq \omega_2$ if

(i) the model set used in the estimation of R_0^f is sufficiently general;

(ii) for some sufficiently large $\mu > 0$, the following condition holds:

$$\frac{\Phi_{v_0^f}(\omega)}{\Phi_{w_0^f}(\omega)} \geq \mu \quad \text{for } \omega_1 \leq \omega \leq \omega_2.$$

It is clear that nothing comes for free, and it is prudent to ask the following question.

- What is the price that we have to pay, in terms of system performance, before a sufficiently high signal-to-noise ratio of the closed-loop output error can be achieved?

We next show that it is necessary to have a certain level of deterioration in robust performance (relative to the effect of noise disturbance) before the closed-loop output error can achieve a sufficiently high signal-to-noise ratio.

By using (12) and (16), we deduce that

$$\frac{\Phi_{v_0^f}(j\omega)}{\Phi_{w_0^f}(j\omega)} \geq \mu \quad \text{for } \omega_1 \leq \omega \leq \omega_2, \quad \omega_1 \geq 0,$$

if and only if

$$|\bar{T}_0^f(j\omega)|^2 > \mu\,|[1 - T_0^f(j\omega)]H(j\omega)|^2 \Phi_e(\omega) \quad \text{for } \omega_1 \leq \omega \leq \omega_2 \quad (17)$$

where $\Phi_e(\omega)$ is the power spectral density of the noise disturbance e. Upon integration, we get

$$\frac{1}{\pi}\int_{\omega_1}^{\omega_2} |\tilde{T}_0^f(j\omega)|^2 \Phi_r(\omega)\, d\omega > \frac{\mu}{\pi}\int_{\omega_1}^{\omega_2} |[1 - T_0^f(j\omega)]H(j\omega)|^2 \Phi_e(\omega)\, d\omega.$$

By Parseval's theorem, and noting that $\omega_1 \geq 0$, we can write

$$J_0^f = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\tilde{T}_0^f(j\omega)|^2 \Phi_r(\omega)\, d\omega,$$

$$J_0^f > \frac{1}{\pi}\int_{\omega_1}^{\omega_2} |\tilde{T}_0^f(j\omega)|^2 \Phi_r(\omega)\, d\omega.$$

Therefore

$$J_0^f > \frac{\mu}{\pi}\int_{\omega_1}^{\omega_2} |[1 - T_0^f(j\omega)]H(j\omega)|^2 \Phi_e(\omega)\, d\omega.$$

We can now restate the conditions necessary for the estimation of R_0^f as follows.

Conclusion 7. We can obtain an unbiased estimate of R_0^f with a small asymptotic normalized variance in a certain frequency range of interest ($\omega_1 \leq \omega \leq \omega_2$; $\omega_1 \geq 0$) if

(i) the model set used in the estimate of R_0^f is sufficiently general;

(ii) there is robust performance deterioration bounded below by

$$\frac{\mu}{\pi}\int_{\omega_1}^{\omega_2} |[1 - T_0^f(j\omega)]H(j\omega)|^2 \Phi_e(\omega)\, d\omega$$

for a sufficiently large $\mu > 0$.

Remark. It is obvious that a problem is ill-posed if the value of σ that specifies the tolerable level of deterioration in robust performance (see Definition 3) does not satisfy the inequality

$$\frac{1}{\pi}\int_{\omega_1}^{\omega_2} |[1 - T_0^f(j\omega)]H(j\omega)|^2 \Phi_e(\omega)\, d\omega < \frac{\sigma}{\mu}$$

for a sufficiently large $\mu > 0$.

4.3. *Practically unbiased estimation of* R_0^f

In Section 4.2 we have shown that the normalized variance of $\hat{R}_0^f$ can be small if the signal-to-noise ratio associated with the closed-loop output error is sufficiently high. However normalized variance can be used as a measure of the quality of an estimate only if the estimate is unbiased. It is therefore necessary to verify that $\hat{R}_0^f$ is a practically unbiased estimate (or an unfalsified model as discussed in Ljung *et al.*, 1991) of R_0^f. In this subsection we show how to infer that $\hat{R}_0^f$ is a practically unbiased estimate of R_0^f by verifying that $(1 + G_1K_0^f)^{-1}G_1K_0^f$ is a practically unbiased estimate of $(1 + GK_0^f)^{-1}GK_0^f$.

We begin by considering

$$\xi_1 = \left(\frac{GK_0^f}{1 + GK_0^f} - \frac{G_1K_0^f}{1 + G_1K_0^f}\right)r + \frac{1}{1 + GK_0^f}He.$$

Clearly, if $(1 + G_1K_0^f)^{-1}G_1K_0^f$ is a practically unbiased estimate of $(1 + GK_0^f)^{-1}GK_0^f$ then the power spectral density of ξ_1 should reflect the effects of the noise disturbance only. We can perform this verification experimentally after G_1 is obtained (as we describe in Section 6). Now recall that, if it is necessary to update the model G_0, the magnitude of

$$\frac{GK_0^f}{1 + GK_0^f} - \frac{G_0K_0^f}{1 + G_0K_0^f} = Q_0^f(1 - \bar{T}_0^f)R_0^f$$

must be significant in a certain frequency range $[\omega_1, \omega_2]$. Therefore, before the model G_0 is updated, both the magnitude of the frequency weighting $Q_0^f(1 - \bar{T}_0^f)$ and the magnitude of R_0^f cannot be small in $[\omega_1, \omega_2]$. Since we can write

$$\frac{GK_0^f}{1 + GK_0^f} - \frac{G_1K_0^f}{1 + G_1K_0^f} = Q_0^f(1 - \bar{T}_0^f)(R_0^f - \hat{R}_0^f),$$

it can easily be deduced that if $(1 + G_1K_0^f)^{-1}G_1K_0^f$ is a practically unbiased estimate of $(1 + GK_0^f)^{-1}GK_0^f$ in $[\omega_1, \omega_2]$ then $\hat{R}_0^f$ is a practically unbiased estimate of R_0^f in $[\omega_1, \omega_2]$. We can therefore conclude the following.

Conclusion 8. We can verify that $\hat{R}_0^f$ is a practically unbiased estimate of R_0^f in $[\omega_1, \omega_2]$ by verifying experimentally that $(1 + G_1K_0^f)^{-1}G_1K_0^f$ is a practically unbiased estimate of $(1 + GK_0^f)^{-1}GK_0^f$ in $[\omega_1, \omega_2]$.

5. MECHANISMS THAT INFLUENCE PERFORMANCE ROBUSTNESS AND IDENTIFICATION

In this section we study mechanisms that influence performance robustness of systems designed by the IMC method. We show that there are three mechanisms that may lead to deterioration in robust performance. However, only one of them contributes to the high signal-to-noise ratio needed for a successful estimation of R_0^f. These observations allow us to deduce situations where the iterative identification and control design process may continue or may terminate prematurely.

Recall that in Section 4.2 we have shown that a certain level of deterioration in robust performance is necessary before we can attempt to find a good estimate of R_0^f. However, we should ask the following question.

- Does it mean that, irrespective of the causes, deterioration in robust performance is always helpful to the identification of R_0^f?

The answer is *obviously no*.

With appropriate substitutions in (16) and (12) respectively, we can obtain

$$v_0^f = \frac{\frac{G-G_0}{G_0}\bar{T}_0^f}{1+\frac{G-G_0}{G_0}\bar{T}_0^f}(1-\bar{T}_0^f)r, \qquad (18)$$

$$w_0^f = \frac{1}{1+\frac{G-G_0}{G_0}\bar{T}_0^f}(1-\bar{T}_0^f)He. \qquad (19)$$

Since $J_0^f = \|v_0^f\|_2^2$, we observe that, disregarding changes in disturbance suppression ability, deterioration in robust performances is governed by the value of

$$J_0^f = \frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_{v_0^f}(\omega)\,\mathrm{d}\omega$$

We therefore conclude from the right-hand side of (18) that for a given reference input, there are three factors that contribute to J_0^f through $\Phi_{v_0^f}(\omega)$.

1. The effect of the term $[(G-G_0)/G_0]\bar{T}_0^f$ in the numerator is independent of the phase angle of $[(G-G_0)/G_0]\bar{T}_0^f$. We call this the *phase-insensitive factor.*
2. The effect of the term $1+[(G-G_0)/G_0]\bar{T}_0^f$ in the denominator depends on the gain and phase margins of $[(G-G_0)/G_0]\bar{T}_0^f$. We call this the *stability margin factor.*
3. The effect of the term $1-\bar{T}_0^f$ depends on the existence of unstable zeros of G_0 within the passband of $\bar{T}_0^f = F_0^f[G_0]_a$. We call this the *unstable-zeros-dependent factor.*

By using (18) and (19), we can write the signal-to-noise ratio associated with the closed-loop output error as

$$\frac{\Phi_{v_0^f}(\omega)}{\Phi_{w_0^f}(\omega)} = \frac{\left|\frac{G(\mathrm{j}\omega)-G_0(\mathrm{j}\omega)}{G_0(\mathrm{j}\omega)}\bar{T}_0^f(\mathrm{j}\omega)\right|^2}{|H(\mathrm{j}\omega)|^2}\frac{\Phi_r(\omega)}{\Phi_e(\omega)}.$$

This equation indicates that for a given reference input and noise disturbance scenario, only an increase in the magnitude of the phase insensitive factor can increase the signal-to-noise ratio of the closed-loop output error. We now summarise the above discussions as follows.

Conclusion 9.

1. There are three factors that can cause the performance robustness to deteriorate: the phase-insensitive factor, the stability margin factor and the unstable-zeros-dependent factor. Among these, only the phase-insensitive factor can contribute to improving the signal-to-noise ratio associated with the closed-loop output error.
2. When the unstable-zeros-dependent factor or the stability margin factor are the main causes of deterioration in robust performance, the signal-to-noise ratio associated with the closed-loop output error may be poor, and it may be difficult to obtain a practically unbiased estimate of R_0^f with a small asymptotic normalized variance. This may cause subsequent difficulties in continuing the iterative identfication and control design process.

Remarks.

- From (18), it is clear that $\Phi_{v_0^f}(\omega)$ cannot be large in the frequency range where the designed sensitivity function has small magnitude. This implies that the frequency range $[\omega_1, \omega_2]$ emphasized in Section 4 cannot be well below λ_0^f.
- From the definitions of the phase-insensitive factor and the stability margin factor, we can deduce that it is possible to estimate R_0^f accurately only in the frequency range where the designed complementary-sensitivity function weighted multiplicative modelling error has large magnitude and small phase lag. This implies that the frequency range $[\omega_1, \omega_2]$ cannot be well above λ_0^f (where $\bar{T}_0^f$ has small magnitude and large phase lag).
- When the stability margin factor or the unstable-zeros-dependent factor are the main causes of deterioration in robust performance, it may be difficult to obtain an accurate estimate of R_0^f. This may lead to premature termination of the iterative identification and control design process. In particular, when the existing model G_0 has unstable zeros within the passband of the designed closed-loop transfer function T_0^f, the designed sensitivity (unstable-zeros-dependent factor) may have large magnitude in a certain frequency region. This fundamental limit in control performance (discussed in Freudenberg and Looze, 1985) causes a deterioration in designed and robust

performances with no improvement in the signal-to-noise ratio associated with the closed-loop output error.

6. IDENTIFICATION AND VALIDATION OF NEW MODELS

In Section 4.2 we have shown that under noisy conditions, the accuracy of the identified model can be improved by increasing the signal-to-noise ratio associated with the closed-loop output error. It was also shown that this is equivalent to having a certain level of deterioration in robust performance relative to the effect of noise disturbance. It is clearly undesirable from the control point of view for robust performance to deteriorate too seriously, while on the other hand it is necessary to have a sufficiently high signal-to-noise ratio in the closed-loop output error before identification can successfully be carried out. Furthermore, it is important to ensure that a model with the right properties is identified. We should therefore like to ask the following practical questions.

1. When should we try to identify a better model?
2. Have we actually identified a good model for our purpose?

Before we can answer these questions, we need methods for validating an identified model. In Section 6.1 we describe a frequency-domain method for model validation. In Section 6.2 we give a time-domain method for model validation. In Section 6.3 we draw on the results of Lee *et al.* (1994), which compared the two methods of model validation and suggest a procedure for identifying a better model.

6.1. *A frequency-domain method for model validation*

In the following we present a model validation method in the frequency domain. It should be emphasized that the model validation procedure is designed with the closed-loop control objective in mind.

Recall that, given the existing model G_0, it is necessary to identify an improved model G_1 when $J_0^f = \|v_0^f\|_2^2$ is excessively large. Evidently ξ_0^f could be large (implying undesirable performance) with one or both of v_0^f and w_0^f large. If the former is larger, there is a potential to reduce it by improved model identification. But this only works (in a particular frequency band $[\omega_1, \omega_2]$) if the signal-to-noise ratio is sufficiently high. Specifically, when only finitely many input–output measurements are available for identifying R_0^f (which parametrizes G), it was shown in Section 4.2 that the normalized variance of $\hat{R}_0^f$ is small only if the signal-to-noise ratio $\Phi_{v_0^f}(\omega)/\Phi_{w_0^f}(\omega)$ associated with $\xi_0^f = v_0^f + w_0^f$ is sufficiently high. Obviously, then one needs to estimate power spectra for w_0 and v_0 (or more precisely ξ_0). We now proceed as follows.

From $v_0 = (T_0 - \bar{T}_0)r$ and $\xi_0 = v_0 + w_0$, we observe that when $r = 0$, the sole contributor to ξ_0 *is* w_0. Therefore we can compute $\Phi_{w_0}(\omega)$ after measuring ξ_0 with $r = 0$. When $r \neq 0$, we have $\xi_0 = v_0 + w_0$. Assuming that v_0 and w_0 are uncorrelated (which follows if r and e are uncorrelated—a typical situation), $\Phi_{\xi_0}(\omega) = \Phi_{v_0}(\omega) + \Phi_{w_0}(\omega)$. By visual comparison of $\Phi_{\xi_0}(\omega)$ with $\Phi_{w_0}(\omega)$, we evaluate the significance of $\Phi_{v_0}(\omega)$ with respect to $\Phi_{w_0}(\omega)$. If $\Phi_{\xi_0^f}(\omega)$ is significantly larger than $\Phi_{w_0^f}(\omega)$ in a frequency band spanning one decade and centred around λ_0^f (when the designed closed-loop bandwidth is λ_0^f), the model G_0 is invalidated for the design of closed-loop systems with bandwidths larger than or equal to λ_0^f.

The method just described can also be used to validate G_1 after it has been identified (both before and after model reduction is performed). We simply replace G_0 by G_1, while retaining K_0^f, in the simulation of the designed closed-loop response to the reference input. This allows us to compute ξ_1 and its power spectrum $\Phi_{\xi_1}(\omega)$. By visually comparing $\Phi_{\xi_1}(\omega)$ with $\Phi_{w_0^f}(\omega)$, we have good confidence that G_1 is a reliable model of G (when the designed closed-loop bandwidth is λ_0^f) if $\Phi_{\xi_1}(\omega)$ is comparable to $\Phi_{w_0^f}(\omega)$ up to λ_0^f.

6.2. *A time-domain method for model validation*

We now describe a time-domain model validation method. This is useful both for establishing that G_0 should be rejected (that is, as a flag for re-identification) as well as for validating a new model, G_1, replacing G_0.

Referring to Fig. 3 and (10), we notice that $e_\beta = L\xi_1$ when $\hat{\Psi}_0^f = 1$, where e_β is the prediction error (also known as the residual). We also observe from (9) that $G_1 = G_0$ when $\hat{R}_0^f = 0$, and from (11) that $\xi_1 = \xi_0^f$ when $G_1 = G_0$. Therefore we have $e_\beta = L\xi_0^f$ when $\hat{\Psi}_0^f = 1$ and $\hat{R}_0^f = 0$. This suggests that G_0 should be rejected if the cross-correlation of the prediction error e_β with the future values of 'input' $\bar{\alpha}$ exceed its (3σ) confidence limits when $\hat{\Psi}_0^f = 1$ and $\hat{R}_0^f = 0$. This reasoning is independent of the true Ψ_0^f. See Ljung (1987) for more details of model validation by correlation techniques. (Actually it is also easy to apply the same method to validate a pair of newly identified $\hat{R}_0^f$ and $\hat{\Psi}_0^f$

before $\hat{R}_0^f$ is used to calculate G_1. We simply check that the correlation of e_β with the future values of $\bar{\alpha}$ are within their respective confidence intervals.)

6.3. *Identification of a better model*

The methods of model validation described in Sections 6.1 and 6.2 were compared critically in Lee *et al.* (1994). The key conclusions of that study are summarized in the following.

- Correlation function estimates and power spectrum estimates are both useful for model validation where the goodness of fit is based on a closed-loop control criterion.
- Correlation function estimates are more sensitive than power spectrum estimates in the sense that the former tend to invalidate a model before identifying a better model is necessary and possible. This *does not imply* that the correlation method is useless. On the contrary, it suggests that the correlation method is useful for detecting incipient modelling errors.
- Power spectrum estimates not only suggest when a model becomes inadequate, but they also indicate the frequency range in which the signal-to-noise ratio is high for identification.
- There is a limit on the achievable accuracy for performing identification on closed-loop systems if existing controllers are designed on the basis of models with unstable zeros. (Recall the effect of the unstable-zeros-dependent factor remarked upon at the end of Section 5.) Therefore prior knowledge of unstable zeros in the existing model (say G_0) is important. Specifically, let ω_z be the minimum critical frequency corresponding to the unstable zeros of G_0; simulation experience confirmed that it is very difficult, if not impossible, to identify a model better than G_0 if $\lambda_0^f \geq \frac{1}{2}\omega_z$. It should be remarked that this is reminiscent of design tradeoffs discussed in Freudenberg and Looze (1985), as opposed to an ill-posed problem.
- In general, we update G_0 if
  (i) both methods of model validation suggest so; and
  (ii) $\lambda_0^f < \frac{1}{2}\omega_z$.

We now suggest a procedure for identifying a better model. Notice that in the frequency range where the current model G_0 has significant modelling errors, the signal-to-noise ratio of the closed-loop output error can be increased by increasing the magnitude of the reference input or by increasing the designed closed-loop bandwidth. If practical operation constraints do not allow the magnitude of the reference input to be increased then the signal-to-noise ratio of the closed-loop output error can only be increased by increasing the designed closed-loop bandwidth. This, however, has the potential danger of causing instability in the actual closed-loop system if the designed closed-loop bandwidth is increased excessively. To avoid this danger, we proceed as follows.

1. Reduce the rate of increasing the designed closed-loop bandwidth λ_0 once the correlation method for model validation has invalidated G_0.
2. Attempt to identify R_0^f (when $\lambda_0 = \lambda_0^f$) as soon as the power spectrum method for model validation suggests that ξ_0^f has a sufficiently high signal-to-noise ratio, provided that $\lambda_0^f < \frac{1}{2}\omega_z$.
   (a) Use the collected data to identify a set of models by experimenting with the likely model structures. Perform model verification on each of these models.
   (b) If an identified model is found to be sufficiently accurate, accept it for the next stage of control design. Otherwise, increase the designed closed-loop bandwidth slightly, collect a new set of measurements and repeat the procedures of model estimation and verification.
   (c) Repeat the last two steps until a sufficiently accurate model is obtained and verified.
3. Terminate the iterative identification and control design procedure if $\lambda_0^f \geq \frac{1}{2}\omega_z$ and ξ_0^f, although unacceptably large, does not facilitate the identfication of a better model.

7. ROBUST PERFORMANCE IMPROVEMENT

Now we know what can be identified and how an identified model can be validated. We have also indicated in Section 3 what we would like to identify. It is therefore logical to ask the following question.

- How does the object which we *can* identify relate to the object which we *would like* to identify?

The answer is that the objects are virtually the same, although it is not obvious. What we can identify is couched in terms of R_0^f, and what we would like to identify is couched in terms of G_1. We need to connect these characterizations. In this section we show that, provided that certain conditions are satisfied, the controller designed on the basis of the model G_1 updated through an estimate of $\hat{R}_0^f$ can improve the performance robustness of the system.

Recall from (17) that, just before we attempt

to update the model G_0 through identifying R_0^f, it is necessary that

$$|\bar{T}_0^f(j\omega)|^2 \Phi_r(\omega) > \mu |[1 - T_0^f(j\omega)]H(j\omega)|^2 \Phi_e(\omega) \quad \text{for } \omega_1 \le \omega \le \omega_2$$

for a sufficiently large $\mu > 0$. Furthermore, it is also necessary that $|\bar{T}_0^f(j\omega)|$ in the above inequality be mainly contributed by the phase-insensitive factor before an accurate estimate of R_0^f can be obtained. This implies that in order to improve the robust performance through identification and redesign, it is necessary that the phase-insensitive factor (which is also the complementary sensitivity weighted multiplicative factor (which is also the complementary sensitivity weighted multiplicative modelling error) associated with the updated model G_1 and the redesigned controller K_1^0 (while keeping $\lambda_1^0 = \lambda_0^f$) be small in the frequency range $[\omega_1, \omega_2]$. Hence it is relevant to consider the magnitude of the ratio

$$\left(\frac{G - G_1}{G_1}\bar{T}_1^0\right) \Big/ \left(\frac{G - G_0}{G_0}\bar{T}_0^f\right)$$

in the frequency range $[\omega_1, \omega_2]$.

Before the main results are presented in Theorems 3 and 4, we state a theorem (which follows directly from Theorem 4 of Lee *et al.*, 1993) that is relevant to the choice of the relative degree of $\hat{R}_0^f$, and establish two lemmas that we use in the proof of Theorem 3.

Theorem 2. Let the controller K_0^f and the proper stable transfer function Q_0^f designed by the IMC method described in Section 2.1. Then the relative degree of

$$R_0^f = \frac{G - G_0}{1 + Q_0^f(G - G_0)}$$

is given by

$$\text{rel deg}\{R_0^f\} = \min(\text{rel deg}\{G\}, \text{rel deg}\{G_0\}).$$

Remark. The relative degree of the strictly proper plant G is usually unknown. It is therefore necessary to allow, in the identification of R_0^f, the relative degree of $\hat{R}_0^f$ to take the smallest possible value of one.

Lemma 1. Suppose that G_0 has a relative degree of $n \ge 1$, $Q_0^f = [G_0]_m^{-1}F_0^f$ is proper, and $\hat{R}_0^f$ has a relative degree of $q \ge 1$. If G_1 is updated according to

$$G_1 = G_0 + \frac{\hat{R}_0^f}{1 - Q_0^f\hat{R}_0^f}$$

then

(i) G_1 has a relative degree k, where

$$k \ge n \qquad \text{if } q = n,$$

$$k = \min(n, q) \quad \text{otherwise;}$$

(ii) there exists a strictly proper IMC filter

$$F_1^0 = F_0^f\left(\frac{\lambda_0^f}{s + \lambda_0^f}\right)^i, \quad i \text{ is an integer}$$

(where F_0^f has a relative degree such that $Q_0^f = [G_0]_m^{-1}F_0^f$ is proper) such that $Q_1^0 = [G_1]_m^{-1}F_1^0$ has at least the relative degree of Q_0^f;

(iii) if G_1 has no zeros along the imaginary axis then Q_1^0/Q_0^f is bounded along the imaginary axis; in particular, there exists a finite δ such that

$$\sup_{\omega_1 \le \omega \le \omega_2} \left|\frac{Q_1^0(j\omega)}{Q_f^0(j\omega)}\right| = \delta.$$

Proof. (i) There are three cases to be considered, namely $q = n$, $q < n$ and $q > n$.

For the case where $q = n$, consider

$$\lim_{s\to\infty} s^q G_1 = \lim_{s\to\infty} s^n G_0 + \lim_{s\to\infty} \frac{s^q\hat{R}_0^f}{1 + q_0^f\hat{R}_0^f}.$$

Since

$$\lim_{s\to\infty} s^n G_0 = c_1, \quad c_1 \ne 0,$$

and

$$\lim_{s\to\infty} \frac{s^q\hat{R}_0^f}{1 + Q_0^f\hat{R}_0^f} = c_2, \quad c_2 \ne 0,$$

we have

$$\lim_{s\to\infty} s^q G_1 = 0 \quad \text{if } c_2 = -c_1,$$

or

$$\lim_{s\to\infty} s^q G_1 \ne 0 \quad \text{if } c_2 \ne -c_1.$$

Hence the relative degree of G_1 is $k \ge q = n$.

Next consider

$$\lim_{s\to\infty} s^q G_1 = \lim_{s\to\infty} s^q G_0 + \lim_{s\to\infty} \frac{s^q\hat{R}_0^f}{1 + Q_0^f\hat{R}_0^f}$$

for $q < n$. Clearly,

$$\lim_{s\to\infty} s^q G_1 = \lim_{s\to\infty} s^q\hat{R}_0^f \ne 0,$$

since Q_0^f is proper and G_0 has a relative degree larger than q. Therefore the relative degree of G_1 is $k = q < n$.

Similar considerations applying to

$$\lim_{s\to\infty} s^n G_1 = \lim_{s\to\infty} s^n G_0 + \lim_{s\to\infty} \frac{s^n \hat{R}_0^f}{1 + Q_0^f \hat{R}_0^f}$$

for $q > n$ will lead to the conclusion that the relative degree of G_1 is $k = n < q$.

(ii) It follows immediately from the above that $Q_1^0 = [G_1]_m^{-1} F_1^0$ has at least the relative degree of $Q_0^f = [G_0]_m^{-1} F_0^f$ if

$$F_1^0 = F_0^f \left(\frac{\lambda_0^f}{s + \lambda_0^f}\right)^i, \quad i \geq k - n.$$

Obviously, since $k \geq 1$ and the relative degree of F_0^f is at least n, F_1^0 is strictly proper.

(iii) Now it is easy to conclude that for the above choice of F_1^0, Q_1^0/Q_0^f is proper. Therefore if G_1 has no zeros along the imaginary axis and G_0 has no poles along the imaginary axis then Q_1^0/Q_0^f is bounded along the imaginary axis and there exists a finite δ such that

$$\sup_{\omega_1 \leq \omega \leq \omega_2} \left|\frac{Q_1^0(j\omega)}{Q_f^0(j\omega)}\right| = \delta.$$

□

Remarks.

- It will be clear from Theorem 3 that it is undesirable for δ to become excessively large.
- Since all poles of G_0 that are also poles of G are always retained by a well-identified G_1, it is clear that poles of G_0 that are also poles of G, even if they are near the imaginary axis, will not cause δ to assume an excessively large value.
- Zeros of G_1 near the imaginary axis for $\omega_1 \leq \omega \leq \omega_2$ may not appear as zeros of G_0. However, these zeros would be zeros of the plant G if G_1 is a well-identified model of G for $\omega_1 \leq \omega \leq \omega_2$. This would happen only if we increased the closed-loop bandwidth to the frequency range where the plant has (stable or unstable) zeros near the imaginary axis, and the controller has *excessively large* gain. Therefore we can prevent δ from being excessively large by observing well-known design guidelines.
- If G_0 has poles near to the imaginary axis for $\omega_1 \leq \omega \leq \omega_2$ that are not poles of the plant G then a well-identified model G_1 for G either will have no poles at these locations or will have approximate pole zero cancellations at these locations. In these situations δ may become excessively large. It is therefore *important to verify* that an identified model (such as G_0) has no unnecessary poles near the imaginary axis.

Lemma 2. If G_0 and G_1 are strictly proper stable models of the plant G, and $\bar{T}_0^f = G_0 Q_0^f$ is the closed-loop transfer function, where Q_0^f is designed by the IMC method, then there exists a finite η such that

$$\sup_{\omega_1 \leq \omega \leq \omega_2} \left|\frac{G_1(j\omega) - G_0(j\omega)}{G_0(j\omega)} \bar{T}_0^f(j\omega)\right| = \eta.$$

Proof. Clearly the transfer function $G_1 - G_0$ is stable. Also, from the facts that $\bar{T}_0^f = G_0 Q_0^f$ and that the Q_0^f designed by the IMC method is proper and stable, it is easy to conclude that

$$\frac{G_1 - G_0}{G_0} \bar{T}_0^f = (G_1 - G_0) Q_0^f$$

is proper and stable. □

Remarks.

- It will be clear from Theorem 3 that it is undesirable for η to become excessively large.
- η may become excessively large if G_1 has poles near the imaginary axis for $\omega_1 \leq \omega \leq \omega_2$ that are not poles of G_0, and if λ_0^f is very near to the critical frequencies of these poles. However, this is impossible because if G_1 were a well-identified model of the plant G then G would have poles near to $\pm j\lambda_0^f$ that are not poles of G_0. Under these conditions, the actual closed-loop system T_0^f would be unstable or almost unstable. Furthermore, λ_0^f cannot be close to the zeros of G_0 near the imaginary axis for $\omega_1 \leq \omega \leq \omega_2$, because this will result in a controller with an excessively large gain in that frequency range. Hence, by ensuring that the actual closed-loop system T_0^f is far from instability (recall the guidelines given at the end of Section 6.3) and by observing well-known controller design guidelines, we automatically prevent η from taking excessively large values.

Theorem 3. Let G_0 be a stable strictly proper model of the plant G. Suppose that G is stabilized by the controller K_0^f designed according to the IMC method described in Section 2.1, and hence has the description

$$G = G_0 + \frac{R_0^f}{1 - Q_0^f R_0^f}, \tag{20}$$

where

$$Q_0^f = [G_0]_m^{-1} F_0^f, \tag{21}$$

with $[G_0]_m$ the minimum-phase factor of G_0. Let

1. When can one redesign the controller and expand the closed-loop bandwidth, without re-identifying?
2. When should one re-identify?
3. What does one want to identify in the re-identification procedure?
4. What can one identify in the re-identification procedure?

In order to check if an identified model is actually good for our purpose, we have presented two methods for validating an identified model experimentally before it is employed in controller redesign.

The main conclusion of this paper is that, given a strictly proper stable model of a strictly proper stable plant, it is possible to improve the robust performance of a closed-loop system through the windsurfer approach if

(i) the deterioration in performance robustness caused by increasing the closed-loop bandwidth is mainly contributed by the phase-insensitive factor;

(ii) the deterioration in performance robustness caused by increasing the closed-loop bandwidth resulted in a sufficiently high signal-to-noise ratio associated with the closed-loop output error;

(iii) the designed closed-loop bandwidth has not approached the minimum critical frequency corresponding to the unstable zeros of the plant or the existing model.

Acknowledgements—The authors wish to acknowledge the funding of the activities of the Co-operative Research Centre for Robust and Adaptive Systems by the Australian Government under the Cooperative Research Centres Program. R. L. Kosut wishes to acknowledge support by AFOSR, Directorate of Mathematical and Computer Sciences, Under Contract F49620-93-C-0012. W. S. Lee wishes to thank Dr Robert R. Bitmead for valuable advice.

REFERENCES

Anderson, B. D. O. and R. L. Kosut (1991). Adaptive robust control: on-line learning. In *Proc. 30th IEEE Conf. on Decision and Control*, Brighton, UK, pp. 297–298.

Campi, M., W. S. Lee and B. D. O. Anderson (1994). New filters for internal model control design. *Int. J. Robust and Nonlinear Control*, **4**, 757–775.

Freudenberg, J. S. and D. P. Looze (1985). Right half plane poles and zeros and design tradeoffs in feedback systems. *IEEE Trans. Autom. Control*, **AC-30**, 555–565.

Gevers, M. (1993). Towards a joint design of identification and control? In H. L. Trentelman and J. C. Willems, (Eds) *Essays on Control: Perspectives in the Theory and its Applications*, pp. 111–151. Birkhäuser. Boston.

Hansen, F. R. (1989). A fractional representation approach to closed-loop system identification and experiment design. PhD dissertation, Stanford University.

Lee, W. S., B. D. O. Anderson, R. L. Kosut and I. M. Y. Mareels (1993). A new approach to adaptive robust control. *Int. J. Adaptive Control and Signal Processing*, **7**, 183–211.

Lee, W. S., B. D. O. Anderson, I. M. Y. Mareels and R. L. Kosut (1994). On some practical issues in system identification for the windsurfer approach to adaptive robust control. In *Proc. 10th IFAC Symp. on System Identification*.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ.

Ljung, L., B. Wahlberg and H. Hjalmarsson (1991). Model quality: the roles of prior knowledge and data information. In *Proc. 30th IEEE Conf. on Decision and Control*, Brighton, UK, pp. 273–278.

Morari, M. and E. Zafiriou (1989). *Robust Process Control*. Prentice-Hall, Englewood Cliffs, NJ.

Partanen, A. G. and R. R. Bitmead (1993). Two stage iterative identification/control design and direct experimental controller refinement. In *Proc. 32nd IEEE Conf. on Decision and Control*, San Antonio, TX, pp. 2833–2838.

Schrama, R. J. P. (1992). Accurate identification for control: The necessity of an iterative scheme. *IEEE Trans. Autom. Control*, **AC-37**, 991–994.

Schrama, R. J. P. and P. M. J. Van den Hof (1992). An iterative scheme for identification and control design based on coprime factorizations. In *Proc. American Control Conf.*, pp. 2842–2846.

Zang, Z., R. R. Bitmead and M. Gevers (1991). Iterative model refinement and control robustness enhancement. In *Proc. 30th IEEE Conf. on Decision and Control*, Brighton, UK, pp. 279–284.

A.5 A family of norms for system identification problems

M. Massoumnia and R.L. Kosut, *IEEE Trans. on Automatic Control*, vol. 39, no. 5, May 1994, pp.1027-1031. Appeared first in *Proc. 1993 ACC*, San Francisco, CA, June 1993.

A Family of Norms for System Identification Problems

Mohammad-Ali Massoumnia and Robert L. Kosut

Abstract— **In this correspondence, we introduce a family of norms that may prove useful in system identification problems. The important property of the new norm is that for a given sequence, its value in the limit will converge to the supremum over all frequencies of the spectrum of the sequence. Using this property, a procedure is outlined to approximately minimize the weighted $\mathcal{L}_\infty$ norm of the frequency response estimation error. In addition, a test for checking the Whiteness of a given sequence is proposed.**

I. Introduction

The parametric approach to system identification is based on selecting an appropriate model structure and a search for the parameters of the model that best describes the data. Usually, the best model within the model set is characterized as the one that minimizes a selected norm of the prediction errors. By far the most popular norm is the sum of the square of the prediction errors—the quadratic norm. In this correspondence, we introduce a new family of norms that seem to be useful in system identification problems. The new norms have interesting interpretation in the frequency domain and include the usual quadratic norm as special case. The important property of the new norm is that in the limit, its minimization is equivalent to minimizing the supremum over all frequencies of the spectrum of the prediction error or equivalently, minimizing its $\mathcal{L}_\infty$ norm.

Manuscript received February 12, 1993. This work was supported by the National Science Foundation under Grant ISI-9161408, and by the AFOSR Director of Mathematics and Information Sciences under Grant F49620-C-0119.

M.-A. Massoumnia was with Integrated Systems, Inc., Santa Clara, CA. He is now with Negar-Afzar, Palo Alto, CA 94309 USA.

R. L. Kosut is with Integrated Systems, Inc., Santa Clara, CA 95054 USA.

IEEE Log Number 9216466.

II. Definitions and Preliminaries

Let us assume we are given a scalar bounded sequence $\{e_i, i = 1, \cdots, N\}$ which in our application represents the prediction errors computed from the observed data and a guessed model parameter vector θ. Based on this sequence, form the $(N+M-1) \times M$ matrix

$$E_{NM} = \frac{1}{\sqrt{N}} \begin{bmatrix} e_1 & 0 & \cdots & 0 \\ e_2 & e_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ e_M & e_{M-1} & \ddots & e_1 \\ \vdots & \vdots & \vdots & \vdots \\ e_N & e_{N-1} & \cdots & e_{N-M+1} \\ 0 & e_N & \cdots & e_{N-M+2} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e_N \end{bmatrix} \tag{1}$$

where we assume $1 \leq M \leq N$. Note that E_{NM} is constant along the diagonals, and for $M = 1$, E_{N1} is a column vector with $e_i/\sqrt{N}$ as its elements. To simplify the notation, we denote this vector by E_N. Moreover, the matrix E_{NM} is completely specified when $E_N (= E_{N1})$ and the value of M are given.

It is simple to see that the matrix $E_{NM}^T E_{NM}$ is symmetric, at least positive semidefinite, and Toeplitz. The elements of this matrix are estimates of the autocorrelation function of the sequence e_i. More explicitly, define the sequence a_i $(i = 0, \cdots, M-1)$ in terms of e_i as

$$a_i = \frac{1}{N} \sum_{j=1}^{N-i} e_j e_{j+i}. \tag{2}$$

Then a simple computation shows

$$E_{NM}^T E_{NM} = \begin{bmatrix} a_0 & a_1 & \cdots & a_{M-1} \\ a_1 & a_0 & \ddots & a_{M-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M-1} & a_{M-2} & \cdots & a_0 \end{bmatrix}. \tag{3}$$

Using these definitions, we define the new norm as the maximum eigenvalue of $E_{NM}^T E_{NM}$

$$V_M(E_N) = \overline{\lambda}(E_{NM}^T E_{NM}) = \overline{\sigma}^2(E_{NM}) \tag{4}$$

where $\overline{\lambda}(F)$ denotes the maximum eigenvalues of F, and $\overline{\sigma}(F)$ denotes the maximum singular value of F. For simplicity, we usually delete the argument of V_M and assume it is understood to be a function of E_N which is itself formed from the prediction errors e_i. Note that V_M defined in (4) is not mathematically a norm on $\mathcal{R}^N$; however, $\sqrt{V_M(E_N)}$ *is a valid norm for* E_N, and only to simplify the presentation we refer to V_M as a norm.

Also, for $M = 1$, V_M is identified with the usual quadratic norm. From another point of view, V_1 only includes an estimate of the autocorrelation function of the prediction error for zero shift, a_0. Moreover, V_M is nicely bounded by V_1 as follows:

$$\|E_N\|_2^2 = V_1(E_N) \leq V_M(E_N) \leq M V_1(E_N) = M\|E_N\|_2^2. \tag{5}$$

To illustrate some of the properties of V_M for $M > 1$, assume $M = 2$. The maximum eigenvalue of $E_{N2}^T E_{N2}$ is simple to compute and is given by

$$V_2 = a_0 + |a_1|. \tag{6}$$

In this case, not only the sum of square of prediction errors is included in the performance measure, but this norm also includes an estimate of the autocorrelation function of the prediction error at the first time shift. Therefore, minimizing V_2 will force $|a_1|$ to small values. This is a first attempt to whitening the prediction error in addition to minimizing its variance.

Note that the whiteness of the prediction error is an important factor in the validating a computed model [1]. this desirable property of the prediction error is not reflected, however, in any form in the usual quadratic norm. But V_M is not only a function of the variance of the prediction error but is also a function of the values of the autocorrelation of the prediction error for time shifts up to $M-1$, and by increasing M more and more of the temporal behavior of this autocorrelation affects V_M.

III. Frequency Domain Properties

Now we discuss the frequency domain interpretation of the new norm. First assume the limit of α_i defined in (2) as N goes to infinity exists

$$\lim_{N \to \infty} a_i = \overline{a}_i. \tag{7}$$

If, in addition, $\overline{a}_i$ is in l_1, then the spectrum of the prediction error is

$$S_{ee}(\omega) = \sum_{k=-\infty}^{\infty} \overline{a}_k e^{-j\omega k} \tag{8}$$

where we set $\overline{a}_{-k} = \overline{a}_k$ because we are dealing with a real sequence. It is shown in [2] that the following are true:

$$\overline{a}_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{ee}(\omega)\, d\omega \tag{9}$$

$$\overline{\lambda}(C_M) \leq \sup_{|\omega| \leq \pi} S_{ee}(\omega) = \lim_{M \to \infty} \overline{\lambda}(C_M) \tag{10}$$

$$\underline{\lambda}(C_M) \geq \inf_{|\omega| \leq \pi} S_{ee}(\omega) = \lim_{M \to \infty} \underline{\lambda}(C_M) \tag{11}$$

where $\underline{\lambda}(F)$ and $\overline{\lambda}(F)$ denote, respectively, the smallest and the largest eigenvalue of F, and the Toeplitz matrix C_M is defined as follows:

$$C_M = \begin{bmatrix} \overline{a}_0 & \overline{a}_0 & \cdots & \overline{a}_{M-1} \\ \overline{a}_1 & \overline{a}_0 & \ddots & \overline{a}_{M-2} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{a}_{M-1} & \overline{a}_{M-2} & \cdots & \overline{a}_0 \end{bmatrix}. \tag{12}$$

Theorem 1: The following limits hold:

$$\lim_{N \to \infty} E_N^T E_N = \frac{1}{2\pi} \int_{\pi}^{\pi} S_{ee}(\omega)\, d\omega \tag{13}$$

$$\lim_{M \to \infty} \left(\lim_{N \to \infty} \overline{\sigma}^2(E_{NM}) \right) = \sup_{|\omega| \leq \pi} S_{ee}(\omega) \tag{14}$$

$$\lim_{M \to \infty} \left(\lim_{N \to \infty} \underline{\sigma}^2(E_{NM}) \right) = \inf_{|\omega| \leq \pi} S_{ee}(\omega) \tag{15}$$

where we assume that N goes to infinity faster than M.

Proof: Relation (13) follows from the definition of $\overline{a}_0$ and (9). Moreover, by definition of C_M, we have

$$\lim_{N \to \infty} E_{NM}^T E_{NM} = C_M. \tag{16}$$

Substituting this in (10) and (11) and noting that the eigenvalues of a matrix are continuous functions of the elements of the matrix, the other results follow immediately. □

In identification problems, we estimate the model parameters θ by minimizing $V_M(E_N(\theta))$. The notation $E_N(\theta)$ emphasizes the fact that the prediction error is a function of θ and the minimization is carried over elements of θ. Relation (14) is very illuminating in this respect and shows that by minimizing V_M as M approaches infinity, the supremum over all frequencies of the spectrum of prediction error is minimized. Because of this property, we refer to the identification problem using the new norm (with large values of M) as the $\mathcal{L}_\infty$

identification problem. In contrast, by minimizing the usual quadratic norm, the integral of the spectrum of prediction error over all frequencies is minimized [1], and this can be referred to as $\mathcal{L}_2$ *identification problem* [see (13)].

Now we explore the usefulness of the new norm in identification problems, and relate the $\mathcal{L}_\infty$ norm of the spectrum of the prediction error to $\mathcal{L}_\infty$ norm of the transfer function estimation error. Following the procedure used in [1], let us assume the true system output is generated by

$$y_t = G_0(q)u_t + v_t \tag{17}$$

where the additive noise v_t has the spectrum

$$S_{vv}(\omega) = \lambda_0 |H_0(e^{j\omega})|^2 \tag{18}$$

with $H_0(\infty) = 1$. Also assume the suggested model for the system has the form

$$y_t = G(q, \theta)u_t + H(q, \theta)e_t \tag{19}$$

where θ is the vector of unknown parameters. Provided u and v are independent, it is simple to show that the spectrum of the prediction error in this case is given by [1]

$$S_{ee}(\omega, \theta) = \frac{|\tilde{G}(e^{j\omega}, \theta)|^2 S_{uu}(\omega) + S_{vv}(\omega)}{|H(e^{j\omega}, \theta)|^2} \tag{20}$$

where $\tilde{G} = G - G_0$ is the error in estimating the transfer function.

Unfortunately, the term $S_{vv}/|H|^2$ in (20), which depends on the parameter θ, prevents us from directly resulting the minimization of S_{ee} to the minimization of $|\tilde{G}|$. Using a procedure similar to the one outlined in [1], we can first use a high-order ARX model

$$A(q)y_t = B(q)u_t + e_t \tag{21}$$

to approximate $H_0(e^{j\omega})$ by $1/A(e^{j\omega})$, and filter both u_t and y_t by $A(q)$. Let us denote the filtered input and output by u_t^f and y_t^f, respectively. Next use the following output error model to estimate the model parameters θ

$$y_t^f = G(q, \theta)u_t^f + e_t. \tag{22}$$

Now using (20) we have

$$S_{ee}(\omega, \theta) = |\tilde{G}(e^{j\omega}, \theta)|^2 |A(e^{j\omega})|^2 S_{uu}(\omega) + |A(e^{j\omega})|^2 S_{vv}(\omega). \tag{23}$$

If $1/|A|$ is a good approximation to $|H_0|$, then the last term in (23) is a constant equal to λ_0, and we can write

$$|\tilde{G}(e^{j\omega}, \theta)|^2 |A(e^{j\omega})|^2 S_{uu}(\omega) \approx S_{ee}(\omega, \theta) - \lambda_0. \tag{24}$$

Using (24), it is clear that *minimizing the supremum of* S_{ee} in this case will directly lead to the minimization of the weighted $\mathcal{L}_\infty$ *norm of* $\tilde{G}$. Note that, as expected, the weighting $|A|^2 S_{uu}$ ($\approx S_{uu}/|H_0|^2$) puts more emphasis on the frequency ranges where the signal-to-noise spectral ratio is large. Also, by repeating the experiment with a different input (changing S_{uu}), we have the flexibility of changing this weighting factor.

Note that if the model used for G in (22) is finite-impulse-response (FIR), then the new norm is a convex function of the parameters; otherwise, the problem in general is nonconvex and may be difficult to solve.

Moreover, after minimizing V_M (for sufficiently large value of M), we can compute a good estimate for the supremum over all frequencies on the left-hand side of (24). Since the supremum over all frequencies of the first term on the right-hand side of (24) can be approximated by the minimum value of V_M, and the value of λ_0 (variance of the noise) can be approximated when we are computing the ARX structure in (21), this gives an estimate for the weighted $\mathcal{L}_\infty$ norm of modeling error.

TABLE I
SAMPLE MEAN, STANDARD DEVIATION, AND 95% CONFIDENCE THRESHOLD OF ζ FOR SEVERAL DIFFERENT VALUES OF M AND N

M	10	20	50
	1.59	3.87	16.77
N=100	0.76	1.92	12.42
	3.08	7.50	36.90
	0.94	2.05	6.88
N = 200	0.36	0.80	3.46
	1.60	3.60	12.30
	0.52	0.99	2.64
N = 500	0.17	0.31	0.84
	0.84	1.53	4.26
	0.34	0.63	1.48
N = 1000	0.11	0.18	0.38
	0.55	0.95	2.15

IV. WHITENESS TEST

Now we digress a little and show how the relations given in (14) and (15) can be used to derive an interesting test for checking the whiteness of a sequence. Note that a white sequence, by definition, has a flat spectrum (impulsive autocorrelation function). Therefore, a sequence is white if and only if the ratio of the supremum to infimum of the spectrum of the sequence is unity. Using this observation, the condition number of E_{NM} as N and M approach infinity (with N faster than M) should converge to unity for a white sequence.

With this introduction, we propose the following condition number test for checking the whiteness of a given sequence. Choose an appropriate value of M and compute the ratio

$$\zeta = \frac{\bar{\sigma}^2 - \underline{\sigma}^2}{\underline{\sigma}^2} \tag{25}$$

where $\underline{\sigma}$ and $\bar{\sigma}$ are the smallest and largest singular values of the E_{NM} matrix. If ζ is less than a specified threshold, then declare the sequence as white; otherwise, declare it as nonwhite. Note that ζ is the condition number of E_{NM} squared minus one, and hence is always positive; and, for a white or slightly correlated sequence, it should be close to zero, and otherwise should be considerably larger than zero.

To use this test, we have to compute the values with which the ratio ζ should be compared for any given values of N and M. Note that assuming we are dealing with a white sequence, it is almost impossible to analytically compute the probability density function of ζ, and from that the necessary 95% confidence threshold. By 95% confidence threshold, we mean the value of α for which the following relation holds:

$$\text{Prob}\,(\zeta \leq \alpha) = 0.95 \tag{26}$$

assuming the sequence under consideration is actually white.

We used a series of simulations to compute these confidence thresholds and the first and second moment properties of ζ assuming the sequence under consideration is actually white. For this purpose, for any given values of N and M, we generated 1000 different random white sequences (of length N), and in each case computed the ratio ζ given in (25). The computed results for different values of M and N are given in Table I. The three numbers in each cell of this table from top to bottom are, respectively, the mean, standard deviation, and 95% confidence threshold of ζ.

Now to recapitulate—we use Table I for checking the whiteness of a sequence as follows. Choose the value of N closest to the length of the sequence. Then choose an appropriate value of M. Usually, M

TABLE II
NUMMBER OF TIMES THE SEQUENCES DID NOT PASS THE WHITENESS TEST AS THE POLE LOCATION r IS VARIED

r	0	0.05	0.1	0.15	0.2
Condition No.	51	115	415	707	940
Chi-Square	54	83	267	577	884

should be chosen between $N/10$ to $N/50$. Next choose a segment of the sequence with length N and form the matrix E_{NM}, and from that compute the ζ given in (25). Next, refer to Table I and check whether the computed value of ζ is less than the required threshold. If this is the case, the sequence is most probably white. Otherwise, it is most probably colored.

We compared the performance of the proposed condition number whiteness test to the chi-square test [1], [3]. We remind the reader that in the chi-square test, the following sum is computed

$$\eta = \frac{N}{a_0^2} \sum_{i=1}^{M-1} a_i^2 \tag{27}$$

where the a_i $(i = 0, \cdots, M-1)$ are defined in (2), and the result is compared to the 95% confidence threshold of the chi-square distribution. The reason for choosing $M-1$ as the upper limit of the summation in (27) is because in the condition number test only the autocorrelation terms up to $M-1$ are present.

For this comparison, we generated (colored) sequences by passing white sequences (which were uniformly distributed at each instant) through the first-order filter

$$H(z) = \frac{1-r}{z-r}. \tag{28}$$

For each pole location (value of r), we filtered 1000 different white sequences (each being 500 points in length) by $H(z)$, and after subtracting the sample mean of the sequence computed the ratio ζ and η, in each case assuming $M = 10$. The number of times the computed ζ was larger than 0.84 (the 95% confidence threshold of condition number whiteness test for $N = 500$ and $M = 10$) is given in Table II. The number of times the computed η was larger than 16.9 (which is the 95% confidence threshold of the chi-square distribution) are also shown in Table II. As can be seen, the condition number test repeatedly performed better in this particular problem [it has a sharper operating characteristic (OC) curve (cf. [4]). We emphasize that the condition number whiteness test is just another test that may prove useful in a particular application.

Moreover, note that for $r = 0$, the output of the first-order filter will be a white sequence, and we expect that about 5% of the 1000 runs (50 runs) will fail the whiteness test, which is very close to the values that are actually obtained in these simulations for both tests.

Finally, we indicate that there are several other whiteness tests that can be devised using the singular values of the E_{NM} matrix. Particularly interesting tests are based on comparing the following ratios with appropriate thresholds

$$\zeta_1 = \frac{\sigma_{av}^2 - \underline{\sigma}^2}{\underline{\sigma}^2} \tag{29}$$

$$\zeta_2 = \frac{\overline{\sigma}^2 - \sigma_{av}^2}{\sigma_{av}^2} \tag{30}$$

with

$$\sigma_{av}^2 = \frac{1}{M} \sum_{i=1}^{M} \sigma_i^2 \tag{31}$$

where σ_i are the singular values of E_{NM}. Using the definition of a_0 and the fact that the trace of the $E_{NM}^T E_{NM}$ matrix in (3) is simply Ma_0, it is clear that we have $a_0 = \sigma_{av}^2$. Moreover, it is clear from (31) that the following inequalities hold

$$\underline{\sigma}^2 \le \sigma_{av}^2 = a_0 \le \overline{\sigma}^2. \tag{32}$$

Note that this relation holds for any finite value of M (with $M \le N$), and is the analog of the following series of inequalities

$$\inf_{|\omega| \le \pi} S_{ee}(\omega) \le \frac{1}{2\pi} \int_{-\pi}^{\pi} S_{ee}(\omega)\, d\omega \le \sup_{|\omega| \le \pi} S_{ee}(\omega). \tag{33}$$

For a white sequence, the left-hand and right-hand side of the inequalities in (32) should converge to the same value, and the tests given in (29) and (30) are based on this observation. Moreover, note that because of (32), the ratios ζ_1 and ζ_2 are always positive, and for a white sequence both converge to zero as M and N approach infinity with N faster than M.

V. CONVERGENCE AND CONVEXITY

The norm introduced in (4) has some interesting properties that we shall discuss next. Let us fix M, and assume we are given a model structure and identify the parameter vector θ of this model by minimizing $V_P(E_N(\theta))$ where P is a positive integer less than M $(P < M)$. Let us assume this optimization problem has a unique global minimum that we will denote by θ^P, and denote the prediction error sequence resulting from this choice of the parameter vector by $E_N^P = E_N(\theta^P)$. Similarly, define θ^M and $E_N^M = E_N(\theta^M)$.

Theorem 2: The following series of inequalities hold:

$$V_P(E_N^P) \le V_P(E_N^M) \le V_M(E_N^M) \le V_M(E_N^P). \tag{34}$$

Proof: Beginning from the left-hand side, the first inequality follows from the fact tat the elements of E_N^P are generated from the model parameters that minimize V_P. The second inequality follows from the interlacing property of the eigenvalues of a symmetric matrix [5]. Note that $(E_{NP}^M)^T E_{NP}^M$ is the first $P \times P$ principle minor of $(E_{NM}^M)^T E_{NM}^M$, where E_{NP}^M and E_{NM}^m are defined in terms of E_N^M using (1). The third inequality follows from the fact that E_N^M is formed from model parameters that minimize V_M. □

The relation given in (34) is especially useful if we set $P = 1$. Then $V_1(E_N^1)$ is the minimum value of the usual least squares performance measure. Also, in this case, we can add another important inequality to the set given in (34).

Corollary 1: The following series of inequalities hold:

$$\begin{aligned} V_1(E_N^1) \le V_1(E_N^m) \le V_M(E_N^M) \le V_M(E_N^1) \\ \le V_1(E_N^1) + (M-1) \max(|a_1^1|, \cdots, |a_{M-1}^2|) \end{aligned} \tag{35}$$

where a_i^1 are computed from the elements of E_N^1 using the relation given in (2).

Proof: The first three inequalities follow by setting $P = 1$ in (34). Moreover, because $Q = (E_{NM}^1)^T E_{NM}^1$ is Toeplitz with a_0^1 on its main diagonal, each eigenvalue of Q denoted by λ satisfies the following inequality:

$$|\lambda - a_0^1| \le (M-1) \max(|a_1^1|, \cdots, |a_{M-1}^1|). \tag{36}$$

This follows from Gershgorin's circle theorem [5], and hence the last inequality in (35) holds. Note that the a_i^1 in (35) are estimates of the autocorrelation functions of the prediction error computed from parameters that are obtained by minimizing the quadratic norm. □

Now let us assume that for a particular problem $V_1(E_N(\theta))$ and $V_M(E_N(\theta))$, both have unique global minimum that are denoted by θ^1 and θ^M, respectively. Moreover, let us assume that in this problem, the last term in (35) goes to zero as the number of data points increases. In other words, assume for a fixed M, we have

$$\lim_{N \to \infty} \max(|a_1^1|, \cdots, |a_{M-1}^1|) = 0. \tag{37}$$

Then, using (35), it is clear that

$$\lim_{N\to\infty} V_1(E_N(\theta^1)) = \lim_{N\to\infty} V_1(E_N(\theta^M)). \tag{38}$$

Now using the assumption on the uniqueness of the global minimum of V_1, it is clear that in the limit θ^1 and θ^M will be identical. Stated more loosely, if the prediction error for the quadratic norm minimization is white, then the parameters obtained by minimizing the new norm will be identical to those obtained by minimizing the usual quadratic norm.

To guarantee that each V_1 and V_M have global minima only, let us choose an ARX model for the structure of the system. In this case, it is well known that the scaled prediction error can be written as

$$E_N = \frac{1}{\sqrt{N}}(Y - \Phi\theta) \tag{39}$$

where Φ is the matrix of regression vectors and Y is the vector of output values [1]. In this case, $V_1(E_N(\theta))$ in a convex function of θ, and the minimization problem has only global minima. Moreover, if Φ is full column rank, then the minimum is unique.

Next we show that *for an ARX model structure*, V_M is also a convex function of the parameters. To see this, note that the matrix E_{NM} can be written as

$$E_{NM} = \sum_{i=1}^{M} T_i E_N w_i^T = \frac{1}{\sqrt{N}} \sum_{i=1}^{M} T_i (Y - \Phi\theta) w_i^T \tag{40}$$

where $w_i \in \mathcal{R}^M$ is the standard basis column vector with 1 in its ith entry and all other elements zero. Also, the $(N + M - 1) \times N$ matrix T_i is defined as

$$T_i = \begin{bmatrix} 0_{(i-1)\times N} \\ I_{N\times N} \\ 0_{(M-i)\times N} \end{bmatrix}. \tag{41}$$

Moreover, denote the ith column of Φ by ϕ_i and the ith element of $\theta \in \mathcal{R}^L$ by θ_i. Then (40) can be rewritten as

$$\begin{aligned} E_{NM} &= C_0 + \sum_{j=1}^{L} C_j \theta_j \\ C_0 &= \frac{1}{\sqrt{N}} \sum_{i=1}^{M} T_i Y w_i^T \\ C_j &= \frac{1}{\sqrt{N}} \sum_{i=1}^{M} T_i \phi_j w_i^T. \end{aligned} \tag{42}$$

Note that $C_j, j = 0, \cdots, L$, have the same special structure as E_{NM}, namely, being constant along the diagonals. Now, using (42), it is clear that E_{NM} is affine in θ, and $\overline{\sigma}(E_{NM})$ is a convex function of θ. Therefore, V_M, being the square of $\overline{\sigma}(E_{NM})$, is also a convex function of θ, and the minimization problem has only global minima in this case.

Using these facts, if we use an ARX model structure, and if it happens that the resulting prediction errors are white (and consequently the relation (37) holds), then we are guaranteed that the parameter estimate using the new norm will be same as the parameters using the quadratic norm for $N \to \infty$. This is promising because, for the sum square norm and ARX structure, there are many established properties [1] that readily extend to the new norm.

If the prediction error sequence is not white, however, which will be the case if the "true" model does not have an ARX structure, then the estimate given by minimizing V_M will usually be different from those obtained from the quadratic norm minimization. Note that the new norm forces the autocorrelation of the prediction error for nonzero shifts to small values (whitens the prediction error), and this property may result in a better estimate of the model parameters (compared to quadratic norm for a given model order) when the true model is not actually inside the model set.

As we have shown previously, for an ARX model structure, the matrix E_{NM} is affine in the parameters, and we are interested in minimizing the maximum singular value of E_{NM}. This problem is already discussed in the literature [6], and a recent algorithm is proposed in [7]. By exploiting the special structure of the matrices C_i defined in (42), however, it may be possible to increase the efficiency of the algorithm in [7]. Also, in our application, the size of the matrices involved is quite large, and special attention should be paid to the memory management and algorithmic implementation: otherwise, huge amounts of memory will be required to perform the optimization even for modest values of M and N.

VI. Conclusion

Although we have presented some preliminary results on the properties of the $\mathcal{L}_\infty$ identification problem in this correspondence, much further work is required to explore the properties of the new norm in details. To perform this task, an efficient implementation of the required minimization algorithm is required so realistic high-order models can be estimated and their properties can be compared to those of the least square minimization. As we previously noted, the convexity of the new norm when an ARX model is used is an important property, and hence many techniques of convex optimization can be used for the solution of this problem.

References

[1] L. Ljung, *System Identification, Theory for the User.* Englewood Cliffs, NJ: Prentice-Hall, 1987.

[2] U. Grenander, and G. Szego, *Toeplitz Forms and Their Applications.* New York: Chelsea, 1958.

[3] G. P. Box, and G. M. Jenkins, *Time Series Analysis, Forecasting and Control.* San Francisco, CA: Holden-Day, 1970.

[4] A. H. Bowker, and G. L. Lieberman, *Engineering Statistics,* 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1972, pp. 158–178.

[5] G. H. Golub, and C. F. VanLoan, *Matrix Computations.* Baltimore, MD: John Hopkins University Press, 1989.

[6] Y. Nesterov, and A. Nemironsky, "Optimization over positive semi-definite matrices: Mathematical background and user's manual," USSR Acad. Sci. Centr. Econ. Math. Inst., 32 Krasikova St., Moscow 117418 U.S.S.R, 1990.

[7] S. Boyd, and L. El Ghaoui, "Methods of centers for minimizing generalized eigenvalues," Stanford Univ., ISL Tech. Rep., 1992.

A.6 Least-squares parameter set estimation for robust control design

R.L. Kosut and B.D.O. Anderson, *Proc. 1994 ACC*, Baltimore, MD, 29 June-1 July, 1994.

Least-Squares Parameter Set Estimation for Robust Control Design

Robert L. Kosut*
Integrated Systems, Inc.
3260 Jay St.
Santa Clara, CA 95054, U.S.A.

Brian D. O. Anderson†
Dept. Of Systems Engineering &
Cooperative Research Center for Robust and Adaptive Systems
The Australian National University
Canberra, ACT 2601, Australia

Abstract Two least-squares based methods are presented for obtaining ARX model sets. The first is obtained using properties of high-order ARX models and the second uses a stochastic embedding scheme on the residuals from an ARX model of any order. Either of the ARX model sets is useful for robust control of systems with uncertain parameters. Using the high order ARX model approach, the parameter uncertainty lies in a confidence ellipsoid. Using the stochastic embedding approach, the parameter uncertainty is ia confidence box. For scalar plants, both cases can be handled using convex programming to obtain the exact stability robustness margin for a particular controller. However, because the uncertainty description is probabilistic, the robustness property has to be associated with a confidence level, *i.e.*, a probability of stability.

1 Introduction

Suppose it is desired to control the single-input-single-output stable *discrete-time* system,

$$y = Gu + He, \quad e \in \mathbf{E}(\lambda) \tag{1}$$

where:

- G and H are known only to be causal linear-time-invariant (LTI) stable systems with unknown transfer functions $G(z)$ and $H(z)$, respectively.
- The sequences y and u are, respectively, the sensed output, and the applied control input. All that is available is the finite data record
$$\{y_t, u_t \mid t = 1, \ldots, N\} \tag{2}$$
where y_t and u_t are the values of the sequences y and u, respectively, at the sample time t.
- The sequence e is unpredictable, but is known to be a member of the set $\mathbf{E}(\lambda)$ where λ is unknown: likely candidiates for $\mathbf{E}(\lambda)$ are $\mathbf{POW}(\lambda)$, the set of sequences with power bound λ, or $\mathbf{IID}(\lambda)$, i.i.d. zero-mean sequences with variance λ.

*Supported by AFOSR, Directorate of Mathematical and Computer Sciences, Contract No. F49620-90-C-0064, NSF U.S./Australia Cooperative Research Program, Travel Grant INT-9014152, and the Australian National University as a Visiting Fellow.

†Supported by the Cooperative Rersearch Center Program of the Australian Government

The problem addressed here, referred to as *set estimation*, is to use the above information to determine a set description of the plant *useful for robust control design*. For example, consider the plant set[1]

$$\mathbf{M} : \left\{ y = (\widehat{G} + \Delta\widehat{W})u + \widehat{H}e \;\middle|\; \|\Delta\|_{\mathcal{H}_\infty} \le \widehat{\gamma},\ e \in \mathbf{E}(\widehat{\lambda}) \right\}$$

If $\mathbf{E}(\widehat{\lambda}) = \mathbf{POW}(\widehat{\lambda})$, then $\mathbf{M}$ is typical starting point for $\mathcal{H}_\infty$ control design. If $\mathbf{E}(\widehat{\lambda}) = \mathbf{IID}(\widehat{\lambda})$, then mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control design methods apply. There are many combinations possible. However, in all the above cases, the quantities with "hats" are available *a priori* to the designer, which is not the case here. These quantities are to be estimated from the data and *a priori* information.

In the remainder of the paper, we show how least-squares estimation can be used to obtain a plant set estimate. We examine two cases: (i) high order model set estimation, and (ii) model set estimation via stochastic embedding. Related techniques can be found in [5] and the references therein.

2 Least-Squares Estimation

Parameter estimation via least-squares with an ARX[2] model structure is perhaps the most widely used approach to system identification. The standard (*e.g.*, [7]) parametric ARX model set is:

$$\mathbf{M}_{\mathrm{ARX}} : \left\{ A_\theta y = B_\theta u + e \;\middle|\; \theta \in \mathbb{R}^p,\ e \in \mathbf{IID}(\lambda) \right\} \tag{3}$$

where

$$A_\theta = 1 + \sum_{i=1}^{n} a_i z^{-i}, \quad B_\theta = \sum_{i=1}^{m} b_i z^{-i}$$

$$\theta = [a_1 \cdots a_n\ b_1 \cdots b_m]^T \in \mathbb{R}^p, \quad p = n + m$$

Thus, the ARX model output at tome t is given by,

$$y_t = \phi_t^T \theta + e_t$$
$$\phi_t^T = [-y_{t-1} \cdots -y_{t-n}\ u_{t-1} \cdots u_{t-m}]$$

[1] If Δ is stable, $\|\Delta\|_{\mathcal{H}_\infty} = \sup_\omega |\Delta(e^{j\omega})|$, otherwize, $\|\Delta\|_{\mathcal{H}_\infty} = \infty$.

[2] ARX is the standard acronymn for system models with an AutoRegressive term (Ay) and an eXternal input (Bu).

The least-squares parameter estimate, based on a finite data record, is found from:

$$\widehat{\theta} = \arg\min_{\theta} \frac{1}{N}\sum_{t=1}^{N}(y_t - \theta^T\phi_t)^2 \tag{4}$$

Limiting Properties It is well known [7] that

$$\widehat{\theta} \to \overline{\theta}, \text{ as } N \to \infty, \text{ w.p. } 1$$

where

$$\overline{\theta} = \arg\min_{\theta} \frac{1}{2\pi}\int_{-\pi}^{\pi} S_{err}(\omega,\theta)d\omega$$

with the "error" spectrum given by,

$$\begin{aligned} S_{err}(\omega,\theta) &= |A_\theta(e^{j\omega})G(e^{j\omega}) - B_\theta(e^{j\omega})|^2 S_{uu}(\omega) \\ &+ \lambda|A_\theta(e^{j\omega})H(e^{j\omega})|^2 \end{aligned}$$

Much more can be said when the system (1) is in the ARX model set (3), *i.e.*, when there exists a parameter $\theta_0 \in \mathbb{R}^p$ such that, $A_{\theta_0}G - B_{\theta_0} = 0$ and $A_{\theta_0}H = 1$. In this case, $\overline{\theta} = \theta_0$ and the parameter error $\widehat{\theta} - \theta_0$ is asymptotically normally distributed *i.e.*, as $N \to \infty$,

$$\sqrt{N}(\widehat{\theta} - \theta_0) \to \mathcal{N}(0, \lambda\mathcal{E}(\phi_t\phi_t^T)^{-1}) \tag{5}$$

where $\mathcal{E}(\cdot)$ denotes expectation. In addition, consistent estimates of $\mathcal{E}(\phi_t\phi_t^T)$ and λ, are given by

$$\widehat{R} = \frac{1}{N}\sum_{t=1}^{N}\phi_t\phi_t^T \quad \widehat{\lambda} = \frac{1}{N-p}\sum_{t=1}^{N}\widehat{e}_t^2$$

with

$$\widehat{e}_t = y_t - \phi_t^T\widehat{\theta}$$

the estimated prediction error.

Large N Properties To simplify notation set

$$\widehat{A} = A_{\widehat{\theta}} \quad \widehat{B} = B_{\widehat{\theta}}$$

Define $\delta \in \mathbb{R}^p$ as the normalized (unknown) parameter error,

$$\delta = \widehat{R}^{1/2}(\widehat{\theta} - \theta_0) \tag{6}$$

Then, the true system (1) can be expressed as:

$$\widehat{A}y = \widehat{B}u - \delta^T\widehat{R}^{-1/2}\phi + e \tag{7}$$

It follows that for sufficiently large N [7]:

$$\delta \in \mathcal{N}(0, \frac{\lambda}{N}) \text{ and } \frac{N}{p}\frac{\delta^T\delta}{\widehat{\lambda}} \in \mathcal{F}(p, N-p) \tag{8}$$

where $\mathcal{F}(p, N-p)$ is the F-distribution with degrees of freedom p and $N-p$. Hence,

$$\text{Prob}\{\delta^T\delta \le \frac{p}{N}\alpha(\eta)\widehat{\lambda}\} = \eta \tag{9}$$

can be determined from an F-distribution table. To be safe, suppose we set η very high, say, $\eta = .999$. Then for typical numbers such as $N \ge 1000$ and $n = m = 10$, we get $\alpha(.999) = 2.27$. For large n, say $n = m = 60$, and large $N \gg n$, we get $\alpha(.999) \approx 1.45$, and so on. In addition, for large N, $e \in \mathbb{E}(\widehat{\lambda})$ with high probability. Hence, for large n, m, and N, the system (1) is in the model set

$$\mathbf{M}_{\text{ARX}} : \begin{cases} \widehat{A}y = \widehat{B}u - \delta^T\widehat{R}^{-1/2}\phi + e \\ \delta^T\delta \le \frac{p}{N}\alpha(\eta)\widehat{\lambda} \\ e \in \mathbb{IID}(\widehat{\lambda}) \end{cases} \tag{10}$$

with probability of at least η. Observe that the model uncertainty in this set is represented as a *norm bound on the normalized parameter error* with a probability tag which can be selected close to one. Putting aside for the moment the issue of whether or not the true system is in the model set, it is therefore logical to pursue robust control with this type of parametric uncertainty.

3 Robust Control with ARX Set

In this section we discuss the issue of robust control design retaining still the assumption that the true system is in the ARX model set $\mathbf{M}_{\text{ARX}}$ of (10). Suppose we apply the LTI feedback controller

$$u = -\widehat{K}y \tag{11}$$

where $\widehat{K}$ stabilizes the "nominal" ARX system ($\delta = 0$),

$$\widehat{A}y = \widehat{B}u + e$$

Applying the control to the actual system model (7), gives the closed-loop system

$$\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} T_\delta \\ -Q_\delta \end{bmatrix} e = \frac{1}{1-\delta^T h}\begin{bmatrix} \widehat{T} \\ -\widehat{Q} \end{bmatrix} e$$

where

$$\begin{aligned} \widehat{T} &= \frac{1}{\widehat{A} + \widehat{B}\widehat{K}}, \quad \widehat{Q} = \frac{\widehat{K}}{\widehat{A} + \widehat{B}\widehat{K}} \\ \widehat{h} &= \widehat{R}^{-1/2}\begin{bmatrix} D\widehat{T} \\ D\widehat{Q} \end{bmatrix}, \quad D = \begin{bmatrix} z^{-1} \\ \vdots \\ z^{-n} \end{bmatrix} \end{aligned}$$

Because $\widehat{K}$ stabilizes the nominal system, $\widehat{T}$, $\widehat{Q}$ and $\widehat{h}$ are all stable.

Recall from the Nyquist theorem that since $\widehat{h}$ is stable, the closed-loop system is stable if and only if, for all ω,

$$|1 - \delta^T\widehat{h}(e^{j\omega})| \ne 0, \quad \forall \delta^T\delta \le r^2 < r_{\text{stab}}^2$$

where r_{stab}, the "real" stability margin is given by,

$$r_{\text{stab}} = \inf_{\omega}\inf_{r(\omega)}\left\{\delta^T\delta = r(\omega)^2 \;\middle|\; \delta^T\widehat{h}(e^{j\omega}) = 1\right\}$$

Calculating $r(\omega)$ involves finding the minimum norm (least-squares) solution to the over-determined set of equations $\delta^T\widehat{h}(e^{j\omega}) = 1$ at each frequency. As shown in [10], an easy calculation yields,

$$r(\omega)^2 = \begin{cases} 1/[\|a\|^2 - (a^Tb)^2/\|b\|^2], & b \ne 0 \\ 1/\|a\|^2, & b = 0 \end{cases} \tag{12}$$

where

$$a = \text{Re } h(e^{j\omega}), \quad b = \text{Im } h(e^{j\omega})$$

Hence, a "probability of stability" can be stated as follows. Since

$$\text{Prob}\{\delta^T\delta \le \frac{p}{N}\alpha(\eta)\widehat{\lambda}\} = \eta$$

it follows that

$$\frac{p}{N}\alpha(\eta)\widehat{\lambda} < r_{\text{stab}}^2 \Rightarrow \text{Prob}\{(1-\delta^Th)^{-1} \text{ stable}\} \ge \eta$$

In the more general case, calculating the stability margin, r_{stab}, for other than the two-norm of the parameter error, can be cast as a a convex programming problem, *e.g.*, [10]. However, no closed form or convex programming solution is known in the MIMO case.

4 High-Order ARX Model Set

The analysis in the previous sections depended on the true system being a member of the parametric ARX model set. Although this is never true, it is a fact (*e.g.*, [7]) that the system (1) is equivalent to the *infinite order* ARX model:

$$Ay = Bu + e \tag{13}$$

where

$$A = H^{-1} = 1 + \sum_{k=1}^{\infty} a_k z^{-k} \quad (14)$$

$$B = H^{-1}G = \sum_{k=1}^{\infty} b_k z^{-k} \quad (15)$$

Because H^{-1} and $H^{-1}G$ are stable, there exists positive constants M_a, M_b, and $\rho < 1$ such that

$$|a_k| \leq M_a \rho^k, \quad |b_k| \leq M_b \rho^k$$

It therefore follows that the n-th order ARX model,

$$A_n y = B_n u + e \quad (16)$$

with

$$A_n = 1 + \sum_{k=1}^{n} a_k z^{-k} \quad (17)$$

$$B_n = \sum_{k=1}^{n} b_k z^{-k} \quad (18)$$

can, for sufficiently large n,arbitrarily well approximate the infinite order ARX model, and hence, the original system. Specifically, it is easily shown that $\|A_n G - B_n\|_{\mathcal{H}_\infty}$ and $\|A_n H - 1\|_{\mathcal{H}_\infty}$ approach zero exponentially as n goes to infinity. As a result, it is very tempting to use the previous result with very high order models, particularly since it is very easy to compute such models. The following example illustrates this approach.

Example The following example is presented without figures – no room! Take the true $G(z)$ as the zero-order-hold equivalent of $\Omega^2/(s^2 + 2\zeta\Omega s + \Omega^2)$ with $\Omega = 2\pi(.1)$ rad/sec, $\zeta = .15$, and a sampling frequency of 2 hz. Take the noise model as $H = 1$ with $e \in \mathbf{IID}(.2^2)$ and with input $u \in \mathbf{IID}(1)$. Hence, the rms noise to signal ratio at the output is rms(e)/rms(Gu) = .28. The system was simulated with u and e drawn from the previous distributions. Least-squares identification was performed for time samples $t = 1 : 512$ for model orders $n = m = 1 : 40$. All the models were validated on a different data set for $t = 513 : 1024$. From the validation set, the minimum rms value occurs at $n = 14$. Because larger orders show an increase in rms, any models of order higher than 14 must be trying to fit the particular noise realization during identification. The fits in magnitude are very similar, but get more noisy at high frequencies. Since the true noise has a constant spectrum, we see that the higher orders are being used primarily to flatten the estimated spectrum. One would therefore expect a significant number of pole-zero cancelations in the estimated transfer function. This phenomena is dramtically visible in a pole zero plot of the estimated transfer function. Moreover, it is interesting to examine the Hankel singular values as well.

An $\mathcal{H}_2$ controller was designed using the 14th order model. (We know this controller stabilizes the true system, because in this case the true system is being simulated. Normally, of course, the true system is unknown to the designer.) Although the closed-loop system is stable and seems to behave well, the robust analysis suggests testing the worst-case plant possible from the ARX model set. By examing the paramaetric stability margins vs. frequency (12) for probability .95 and .995, respectively, it can be seen that at about .5 hz, the closed-loop system is just stable for a .95 probability and just unstable for .995. The interpretion is that there exists a plant in the ARX set which *could* have produced the data with the same statistics. Since this plant would cause considerable trouble for this controller, the controller should be redesigned. , *e.g.*, the loop gain should be reduced in the .5 hz range.

Caveat Emptor There are several impediments to using the high-order ARX estimation and robust control design approach just described. First, if $\mathcal{H}_2$ or $\mathcal{H}_\infty$ methods are useed for design, the controller (11) will also be of high order – although in most cases it is easy to reduce the controller because of the many stable near pole zero cancelations in the estimated model. Secondly, a determination of what is meant precisely by high order is dependent on *a priori* knowledge about the true system. Thirdly, the statistical properties are based on very large data lengths, and a precise value of "large" depends on the true system properties.

To offset the high order problem, an alternative is to use a more parsimonious model parametrization. For example, use of Laguerre or Kautz expansions, as proposed in [11], can result in considerably fewer parameters to obtain the same level of approximation as a model expanded in the backward shift operator z^{-1}. However, the efficacy of this approach depends on prior information regarding the accuracy of some dominant pole locations.

5 Residual Analysis and Stochastic Embedding

The stochastic embedding principle developed by Goodwin *et al.*[3] provides a completely stochastic framework for model error estimation. No assumptions on model order or data length are required. The basic idea is to view model error as a realization of a random variable with zero mean, whose variance is described by a few parameters which captures the structure of the model error. In other words, the complicated problem of relating the bias to the data and the mismatch in structure of the true system and the model is avoided. By modeling the error in this seemingly rudimentary form, an error model set parametrization is obtained which is described by a small number of parameters, yet is capable of representing a large set of error models. In [3] a Maximum Likelihood (ML) approach is used to define the estimates of these model error parameters. In this section, we use the stochastic embedding principle, but applied in a slightly different way, avoiding some of the approximations involved in solving for the ML estimate. Although no prior assumptions about high model order are required, the results presented here – unlike those in [3] – are asymptotic in data length.

To apply the stochastic embedding principle of [3], we proceed as follows: suppose we have obtained ARX estimates $\widehat{A}, \widehat{B}$ from some finite data record. Since the true system (1) is $y = Gu + He$, it follows that the prediction error is:

$$\widehat{e} = \widehat{A}y - \widehat{B}u = \Delta u + w \quad (19)$$

with

$$\Delta = \widehat{A}G - \widehat{B} \quad (20)$$

$$w = \widehat{A}He \quad (21)$$

As before, our goal is to make some kind of estimate covering Δ. Towards this end, form the (asymptotic) prediction error correlation coefficients:

$$c_k = \mathcal{E}(\widehat{e}_t \widehat{e}_{t+k}), \quad \forall k \geq 0 \quad (22)$$

If $u \in \mathbf{IID}(\lambda_u)$, then:

$$c_k = \begin{cases} \lambda_w + \lambda_u \sum_{i=1}^{\infty} \delta_i^2, & k = 0 \\ \lambda_u \sum_{i=1}^{\infty} \delta_i \delta_{i+k}, & k \geq 1 \end{cases} \quad (23)$$

where

$$\lambda_w = \mathcal{E}(w_t)^2 \quad (24)$$

and $\{\delta_1, \delta_2, \ldots\}$ are the impulse response coefficients of Δ, *i.e.*,

$$\Delta(z) = \sum_{i=1}^{\infty} \delta_i z^{-i} \tag{25}$$

Since Δ is stable, there exist positive constants M_δ and $\rho < 1$ such that

$$|\delta_i| \leq M_\delta \rho^{i-1} \tag{26}$$

Hence,

$$|c_k| \leq \begin{cases} \lambda_w + \lambda_u \dfrac{M_\delta^2}{1-\rho^2}, & k = 0 \\ \lambda_u \dfrac{\rho^k M_\delta^2}{1-\rho^2} & k \geq 1 \end{cases} \tag{27}$$

Observe that for $k \geq 1$, the correlation coefficients decay exponentially with k. Hence, information is contained in the correlation coefficients about the model error Δ in the form of bounds on the impulse reponse coefficients $\{\delta_1, \delta_2, \ldots\}$. The problem with (27) is that only bounds are provided and thus it is difficult to infer estimates of M_δ and ρ from c_k or estimates of c_k. The key is to utilize the stochastic embedding put forth in [3].

Stochastic Embedding Make the further assumption that the model error is a realization of a stochastic process such that:

$$\frac{\delta_i}{\rho^{i-1}} \in \mathbf{IID}(\lambda_\delta) \tag{28}$$

As a result, there holds

$$\mathcal{E}_\delta c_k = \begin{cases} \lambda_w + \lambda_u \dfrac{M_\delta^2}{1-\rho^2}, & k = 0 \\ 0, & k \geq 1 \end{cases} \tag{29}$$

Since there is no information in the mean regarding the pulse reponse decay, it is necessary to compute the variance. Thus,

$$\mathcal{E}_\delta c_k^2 = \lambda_u^2 \frac{\lambda_\delta^2 \rho^{2k}}{1-\rho^4}, \quad \forall k \geq 1 \tag{30}$$

We restrict attention now to $k \geq 1$ becasue the information we seek appears only in those coefficients. Observe that the model error has been captured by a parametric stochastic description with just two free parameters λ_δ and ρ.

There are now several ways to fit the sample variance to this expression and find values for λ_δ and ρ which achieve the best fit. One approach to estimating these parameters is to fit a gaussian distribution to the correlation coefficients $c_k, k \geq 1$. That is, assume that $\{c_k \mid k = 1 : L\}$ is an independent gaussian sequence with zero mean and variance $\alpha\beta^k$ where $0 < \beta < 1$. Thus,

$$\beta = \rho^2, \quad \alpha = \lambda_u^2 \frac{\lambda_\delta^2}{1-\rho^4}, \tag{31}$$

and estimates for λ_δ and ρ can be obtained from estimates of α and β. That is, given α and β,

$$\rho = \sqrt{\beta}, \quad M_\delta = \sqrt{\lambda_\delta} = \left(\frac{\alpha(1-\beta^2)}{\lambda_u^2} \right)^{1/4} \tag{32}$$

Taking a max-likelihood approach, the negative log likelihood function is then:

$$\mathcal{L} = \frac{1}{2\alpha} \sum_{k=1}^{L} c_k^2 \beta^{-k} + \frac{L}{2} \log \alpha + \frac{L(L+1)}{4} \log \beta + \text{ constant} \tag{33}$$

The minimizing values $\widehat{\alpha}$ and $\widehat{\beta}$ satisfy:

$$\widehat{\alpha} = \frac{1}{L} \sum_{k=1}^{L} c_k^2 \widehat{\beta}^{-k} \tag{34}$$

$$0 = \sum_{k=1}^{L} \left(k - \frac{L+1}{2} \right) c_k^2 \widehat{\beta}^{-k} \tag{35}$$

Clearly a solution exists iff there is a real root $0 < \widehat{\beta} < 1$. By replacing c_k^2 with $\mathcal{E}_\delta(c_k^2)$ we get the polynomial in $\beta/\widehat{\beta}$:

$$0 = \sum_{k=1}^{L} \left(k - \frac{L+1}{2} \right) \left(\frac{\beta}{\widehat{\beta}} \right)^k$$

Clearly $\beta = \widehat{\beta}$ is a solution, and certainly the desired one in that $\widehat{\alpha} = \alpha$. In fact, it can be shown that, on average, the only real solution is $\beta = \widehat{\beta}$. Even though L can be quite large, *e.g.*, $L = 1024$ is typical, it is actually very easy to solve for $\widehat{\beta}$ by a simple bisection search over $\widehat{\beta}$. Although this method is similar to that developed in [3], here an exact solution to the max-likelihood problem is possible because the data is taken to be the asymptotic prediction error correlation coefficents, *i.e.*, *infinite* data. Clearly errors will be introduced in computing the coefficients using *finite length* data. This requires further analysis.

Having found the estimate $\widehat{\rho}, \widehat{\lambda}_\delta$, an estimate for λ_w follows from the expression for c_0, that is:

$$\widehat{\lambda}_w = c_0 - \lambda_u \frac{\widehat{\lambda}_\delta}{1-\widehat{\rho^2}} \tag{36}$$

Hence, the end result of the max-likelihood/stochastic embedding approach is the following ARX model set:

$$\mathbf{M}'_{\mathrm{ARX}} : \begin{cases} \widehat{A}y = \widehat{B}u + \Delta u + w \\ \Delta = \displaystyle\sum_{k=1}^{\infty} q_k \widehat{\rho}^{k-1} z^{-k} \\ q_k \in \mathbf{IID}(\widehat{\lambda}_\delta) \\ \mathcal{E}(w_t)^2 = \widehat{\lambda}_w \end{cases} \tag{37}$$

6 Robust Control Analysis

In this section we discuss the issue of robust control design under the assumption that the true system is in the ARX model set $\mathbf{M}'_{\mathrm{ARX}}$ of (37). Following the discussion in section , the control $u = -\widehat{K}y$ which stabilizes the nominal ARX system $\widehat{A}y = \widehat{B}u + e$, will stabilize plants in $\mathbf{M}'_{\mathrm{ARX}}$ iff the closed-loop transfer function

$$\frac{1}{1 - \sum_{k=1}^{\infty} \widehat{h}_k(z) q_k} \tag{38}$$

is stable where $\widehat{Q} = \widehat{K}/(1 + \widehat{G}\widehat{K})$ as before and now

$$\widehat{h}_k(z) = \widehat{Q}(z) \widehat{\rho}^{k-1} z^{-k} \tag{39}$$

Because $q_k \in \mathbf{IID}(\widehat{\lambda}_\delta)$, stability is meant again as a "probability of stability." In this case, the uncertain parameters q_k are bounded in magnitude, rather than in an ellipsoid as before. Specifically, since

$$\mathrm{Prob}\{\sup_k q_k^2 \leq \alpha(\eta)\widehat{\lambda}\} = \eta$$

it follows that

$$\alpha(\eta)\widehat{\lambda} < r_{\mathrm{stab}}^2 \Rightarrow \mathrm{Prob}\{(1 - \sum_{k=1}^{\infty} \widehat{h}_k(z) q_k)^{-1} \text{ stable}\} \geq \eta$$

In this case the stability margin is defined as:

$$r_{\mathrm{stab}} = \inf_{\omega} \inf_{r(\omega)} \left\{ q_k^2 = r(\omega)^2 \;\middle|\; \sum_{k=1}^{\infty} \widehat{h}_k(e^{j\omega}) q_k = 1 \right\} \tag{40}$$

Although this is an infinite dimensional convex programming problem, nonetheless a solution is readily obtained *e.g.*, [10].

7 Concluding Discussion

Two least-squares based methods have been presented for obtaining ARX model sets. The first is obtained using properties of high-order ARX models and the second uses a stochastic embedding scheme on the residuals from an ARX model of any order. Various asymptotic approximations are used in obtaining either set, and their effect is under study. Either of the ARX model sets is useful for robust control of systems with uncertain paramaters. However, in this case the uncertainty is random and so a notion of probability of stability is introduced. In other unpublished work of the authors, methods are being developed for analyzing the more practical situation where the data set is finite.

In this paper only some of the many properties of least-squares parameter estimation with ARX models have been examined. There are many properties of least-squares estimation that may yet be profitable for obtaining set estimators useful for robust control design. For example:

1. The least-squares solution is unique and provides the global minimum in closed-form, which in the ideal case allows for a complete asymptotic statistical analysis.
2. By utilizing "square-root" numerical algorithms, *i.e.*, SVD or QR algorithms, the solution can be efficiently and rapidly computed even for very high orders. In particular, with the QR method:
   (a) all models are simultaneously computed up to a specified maximum order and stored in a square-root matrix whose size is determined by the maximum model order, not data length.
   (b) prediction error variance vs. model order (up to the maximum) is available as the last "column" of QR.
   (c) results of different experiments are easily joined, *i.e.*, there is no need to re-compute over both data sets.
   (d) high-order models are easily computed; very rapid computation is possible with lattice algorithms.
3. The impulse response coefficients, up to the order of the numerator polynomial, are asymptotically unbiased.
4. The estimates have orthogonality properties which are not dependent on assumptions about the character of the uncertain disturbance, *e.g.*, gausssian, worst-case, and so on.

Propery 3 is quite appealing. Specifically, if the input u is white, then the first m impulse response coefficients of the true input/output transfer function G are asymptotically unbiased. Recall that m is the order of B_θ, the numerator polynomial in the ARX model set. It is interesting that the result does not depend on n, the order of A_θ, the ARX denominator polynomial. Moreover, the true system need not be in the ARX model set!

This result was originaly proven in [9]. Other useful results following from this fact can be found in [1]. To summarize the result, let

$$\overline{\theta} = \arg\min_{\theta\in\mathbb{R}^p} \mathcal{E}(y_t - \theta^T\phi_t)^2 \tag{41}$$

$$\overline{G} = \frac{\overline{B}}{\overline{A}} = \sum_{k=1}^{\infty} \overline{g}_k z^{-k} \tag{42}$$

Thus $\{\overline{g}_1, \overline{g}_2, \ldots\}$ is the pulse response sequence associated with the asymptotic transfer function estimate of G in (1) which has the pulse response sequence $\{g_1, g_2, \ldots\}$. That is,

$$G = \sum_{k=1}^{\infty} g_k z^{-k}$$

The result from [9] is as follows:

$$u \in \mathrm{IID}(\lambda_u) \Longrightarrow \overline{g}_k = g_k, \quad \forall k = 1:m \tag{43}$$

This result together with the use of an appropriate affine model expansion (*e.g.*, Laguerre or Kautz) may prove most beneficial. An affine model set, *e.g.*, a Laguerre expansion for G, can offset the issue of determining what is meant by a large data length. Moreover, with this model, it is possible to precisely compute statistical properties for any given model order or data length – no asymptotic assumptions are required, *e.g.*, [6]. The stochastic embedding scheme developed in [3] uses such an approach, and if applied as proposed here, it may be possible to eliminate all the asymptotic requirements, *i.e.*, large data length or high model order.

References

[1] H. Aling and R.L. Kosut, "Unbiased least-squares estimates with structure incompatibilities," *Proc. 1993 ECC*, Groningen, The Netherlands, 29 June-2July, 1993.

[2] B. D. O. Anderson *et al.*, *Stability of Adaptive Systems: Passivity and Averaging Analysis*, MIT Press, Cambridge, MA, 1986.

[3] G.C. Goodwin, M. Gevers, and B. Ninness, "Quantifying the error in estimated transfer functions with applications to model order selection," *IEEE Trans. on Automatic Control*, vol. 37, no.7, pp. 913-928, July 1992.

[4] Ulf Grenander and Gabor Szego, *Toeplitz Forms and Their Applications*, Chelsea Pub. Co. 1st Edition 1958, 2nd Edition 1984.

[5] "Special Issue on System Identification for Robust Control Design," edited by R.L. Kosut, G.C. Goodwin, and M.P. Polis, *IEEE Trans. on Automatic Control*, vol. 37, no.7, pp. 899-1008, July 1992.

[6] R.L. Kosut and B.D.O. Anderson, "Statistical analysis of least-squares identification for robust control design: output error case with affine parametrization," *Proc. 1993 ACC*, San Francisco, CA, June 1993.

[7] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, Inc., 1987.

[8] M. Massoumnia and R.L. Kosut, "A family of norms for system identification problems," *Proc. 1993 ACC*, San Francisco, CA, June 1993.

[9] C.T. Mullis and R.A. Roberts, "The use of second-order information in the approximation of discrete-time linear systems, *IEEE Trans. Acoustics, Speech, and Sig. Proc.*, vol. ASSP-24, No.3, pp.226-238, June 1976.

[10] A. Rantzer, "Convex robustness specifications for real parametric uncertainty in linear systems," *Proc. 1992 ACC*, pp. 583-587, June 1992.

[11] B. Wahlberg and L.Ljung, "Hard frequency-domain model error bounds from least-squares like identification techniques", *IEEE Trans. Aut. Contr.*, vol. 37, no. 7, July, 1992.

A.7 Uncertainty model unfalsification: a new system identification paradigm compatible with robust control design

R. L. Kosut, *Proc. 1995 IEEE CDC*, New Orleans, LA, Dec. 1995.

Proc. 1995 CDC
New Orleans, LA
Dec. 1995

Uncertainty Model Unfalsification: A System Identification Paradigm Compatible with Robust Control Design

Robert L. Kosut*†

Abstract To falsify or unfalsify – that is the question.

It is shown that unfalsification of the standard robust control design uncertainty model is a natural replacement for system identification when the intended use of the model is robust control design. For the ARX model, the unfalsification step requires solving a set of convex programming problems, specifically LMI problems, of which ordinary least-squares is one member. The result is a tradeoff curve between model uncertainty and disturbance uncertainty. Hence, a family of models are unfalsified from the data record. The tradeoff curve is given a frequency domain interpretation via the DFT and related computational issues are discussed.

Introduction

Traditional system identification must be modified in order to be compatible with robust control design. The nature of the modification arises from two fundamental principals underlying robust control design: (1) models are inaccurate, and (2) a good control design can withstand model inaccuracy. Hence, robust control design does not require a unique model. But traditional identification extracts a unique model from the data, *i.e.*, as described by Ljung [11], a criterion is posed and minimized over a set of candidate models. Unfortunately, there is no means to provide a measure of the identified model uncertainty, *i.e.*, bounds on model error and disturbances.

In contrast to system identification, a model is said to be *validated* if and only if it could have produced the data. Validation is perhaps a misnomer, as one can never *prove* that a model will be able to accurately predict the future. However, the data can *falsify* a model, *i.e.*, the model may prove to be incapable of producing the data. Hence, a precise term is *unfalsification.*

Clearly unfalsification is a **feasibility problem** – find a member of the model set that is consistent with the data. Identification poses an **optimization problem**, *e.g.*, pick the unfalsified model that has the smallest noise variance. Why? It is certainly possible that some other unfalsified model would produce better closed-loop performance.

In consequence, for robust control design, it is logical to replace system identification with unfalsification. In addition, postulate a model set suitable for robust control, *i.e.*, an *uncertainty model* which includes both disturbance and dynamic uncertainty descriptors. Thus, in summary, two changes must be made in the system identification process:

1. postulate a candidate model set which includes descriptors for both disturbance and dynamic uncertainty
2. use the data to produce all unfalsified models in the candidate model set.

*Author is with Integrated Systems, Inc., 3260 Jay St., Santa Clara, CA 95054, kosut@isi.com

†Research supported by the Air Force Office of Scientific Research, Directorate of Mathematical & Computer Sciences, under Contract No. F49620-93-C-0012.

Having produced a set of *unfalsified uncertainty models*, the remaining question is how to select an appropriate model. Since the ultimate goal is closed-loop design, we are naturally led to evaluating each unfalsified uncertainty model by designing and implementing a robust controller based on the model. Thus, we have the following iterative scheme:

1. Find *all* the unfalsified uncertainty models corresponding to the current data set.
2. Test each unfalsified model, in some order, by applying a robust controller based on a particular unfalsified uncertainty model.
3. Those controllers that yield performance better than or equal to that predicted do not falsify the underlying design model. As soon as performance is worse than predicted, the underlying model is falsified. The unfalsification process can then be repeated with this new data.

Hence, actual measured data is used to rank the models via the effect of a corresponding robust control design.

This paper principally focuses on the unfalsification step with a brief discussion of a particular iterative scheme which follows the broad outline described above.

Previous Research Several approaches have been put forward which involve iterating between identification and control design, *e.g.*, [15, 8, 20, 7]. A good survey is provided by Gevers [4]. A fundamental difficulty with iterative approaches is that in order to prove (analytically) that the iteration will work requires solving very difficult fixed-point problems, *e.g.*, [1].

The framework for uncertainty model unfalsification and robust iterative control design presented here, which avoids some of the difficulties encountered with many iterative schemes, was inspired by the work of Poolla *et al.* [13]. This approach is also heralded in the short essay by Dahleh and Doyle [3], with further development in the paper by Livestone *et al.* [9] and dissertation [10]. The uncertainty tradeoff curves derived here was first presented in [7], but at that time there was no framework upon which to provide a *raison d'etre* for the result.

These approaches differ from that proposed by Safonov *et al.* [17, 18, 14] where a controller set is postualted and members are unfalsified directly from input/output data. In a sense, no model is required except that implied by the performance criterion. In addition, controllers can be falsified without having to be implemented. In model unfalsification as described here, a natural ordering of models is provided based on model error magnitude. Thus, a combination of model and controller falsification may prove to be the most practical.

Organization of Paper The first section provides a brief review of classical least-squares prediction error methods of

system identification. This motivates uncertainty modeling and the subsequent unfalsification procedure which links identification and robust control in a natural iteration. Frequency domain interpretations are presented which provide a means for efficient computation.

Prediction Error (PE) Identification

This section provides a brief review of the classical parametric prediction error method of system identification. The formulation and notation closely follows the presentation in the textbook by Ljung [11].

Data and True System Let u^N and y^N denote, repectively, input and output data which has been recorded at the N integer sample times $t = 1 : N$. Thus,

$$y^N = \{y(t)\}_{t=1}^N, \quad u^N = \{u(t)\}_{t=1}^N \tag{1}$$

For purposes of analysis, we make the following assumptions about the true, but unknown system, which generated the data:

- The true system is stable and operating in open loop, with input/output behavior described by,

$$y = Gu + v \tag{2}$$

  where G is a causal, stable, linear-time-invariant system with unknown (zero-order-hold) transfer function $G(z)$, and v is a disturbance with spectrum $S_{vv}(\omega)$ which is independent of u, *i.e.*, $S_{uv}(\omega) = 0$. It is also assumed that u has spectrum $S_{uu}(\omega)$.

The identification problem is to map the finite data (1) into estimates of $G(z)$ and $S_{vv}(\omega)$, with the additional constraint that we want the estimates to be "good enough" for closed-loop control design. The first step is to pick a candidate model set which might be able to explain the data.

Prediction Error (PE) Model Classical prediction error identification postulates the model set:

$$y = G_\theta u + H_\theta w, \quad \theta \in \Theta, \quad w \in \mathbf{W}_{\mathrm{wht}}(\sigma) \tag{3}$$

where w is an unpredictable sequence drawn from the set $\mathbf{W}_{\mathrm{wht}}(\sigma)$ consisting of white noise sequences with variance σ^2, G_θ and H_θ are causal, linear-time-invariant systems, initially at rest, with transfer functions $G_\theta(z)$ and $H_\theta(z)$, respectively, and are dependent on a parameter vector θ in the set,

$$\Theta = \left\{\theta \in \mathrm{R}^p \;\middle|\; H_\theta^{-1} \text{ and } H_\theta^{-1}G_\theta \text{ are stable}\right\} \tag{4}$$

As explained in [11], parameters in the above set insure that the *predictor* associated with the model set is stable. Specifically, the corresponding one-step ahead prediction error is,

$$e_\theta = H_\theta^{-1}(y - G_\theta u) \tag{5}$$

Hence, $\theta \in \Theta$ implies that $(y, u) \rightarrow e_\theta$ is a stable system. Moreover, if the data (1) was actually generated from (3), then $e_\theta(t) \rightarrow w(t)$ as $t \rightarrow \infty$.

It is often the case that the prediction error is modified by a *data filter*, $F(z)$, that is, e_θ is replaced by Fe_θ. Typically, F is a (stable and minimum phase) pass-band filter. As will be seen, the data filter is particularly important for indirectly manipulating model error (bias).

Observe that the data filter could be considered part of the disturbance model by replacing H_θ with $F^{-1}H_\theta$. Hence, an alternative form of the model set (3) is:

$$y = G_\theta u + F^{-1}H_\theta w, \quad \theta \in \Theta, \quad w \in \mathbf{W}_{\mathrm{wht}}(\sigma) \tag{6}$$

which leads to the (filtered) prediction error

$$e_\theta = FH_\theta^{-1}(y - G_\theta u) \tag{7}$$

The filtered model (6) will be used throughout the remainder of the paper and referred to as the *prediction error model*, abbreviated by PE model.

Least-Squares Identification System identification, in the prediction error context, usually means performing the following "least-squares" optimization:

$$\theta_N = \arg \min_{\theta \in \Theta} \frac{1}{N} \sum_{t=1}^{N} e_\theta^N(t)^2 \tag{8}$$

where $e_\theta(t)$ is an element of the prediction error data sequence,

$$e_\theta^N = \{e_\theta(t)\}_{t=1}^N = FH_\theta^{-1}(y^N - G_\theta u^N) \tag{9}$$

When the true system is in the postulated model set, the optimal error spectrum is constant and equal to the variance of the white noise sequence. When the true system is not in the model set – the usual case in practice –there is no true parameter and the estimated transfer function has a "biased" frequency response, *i.e.*, there is no $\theta \in \Theta$ such that $G(e^{j\omega}) - G_\theta(e^{j\omega}) = 0$ for all ω

As shown in [11], the mean-square model error can be computed asymptotically for large data length N and large model order p. Specifically, for large N, large p, and small p/N,

$$\mathbf{E}\left|G(e^{j\omega}) - G_{\theta_N}(e^{j\omega})\right|^2 \sim \frac{p}{N} \frac{S_{vv}(\omega)}{S_{uu}(\omega)}$$

where $\mathbf{E}$ denotes expectation with respect to the noise v in (2). However, a quantification of "large" or "small" is not known and thus the above expression cannot be used to reliably predict model error.

ARX Model Structure A practical difficulty with using the general model structure (6), is that there is no guaranty that the least-squares criterion has a unique global minimum, although in most cases standard descent methods of optimization will attain a local minimum. Partly because of this reason, the most widely used model structure is the ARX model where:

$$\begin{aligned} G_\theta(z) &= B_\theta(z)/A_\theta(z), \quad H_\theta(z) = 1/A_\theta(z) \\ A_\theta(z) &= 1 + a_1 z^{-1} + \cdots + a_n z^{-n_A} \\ B_\theta(z) &= b_1 z^{-1} + \cdots + b_n z^{-n_B} \\ \theta &= [a_1 \cdots a_{n_A}\ b_1 \cdots b_{n_B}]^T \in \mathrm{R}^p, \quad p = n_A + n_B \end{aligned}$$

Hence, the standard ARX model is:

$$A_\theta y = B_\theta u + F^{-1}w, \quad \theta \in \mathrm{R}^p, \quad w \in \mathbf{W}_{\mathrm{wht}}(\sigma) \tag{10}$$

with the corresponding prediction error,

$$e_\theta = F(A_\theta y - B_\theta u) \tag{11}$$

which is affine in θ. Thus, least squares optimization is convex, and the minimizing θ is readily and efficiently obtained numerically using SVD, QR, or lattice methods.

Modeling for Robust Control

The type of model used for robust control design, referred to as an *uncertainty model* incorporates an additional term which accounts for a class of dynamic uncertainty.

PE Uncertainty Model Specifically, the PE model (6) is modified as follows to obtain the uncertainty model:

$$y = G_\theta u + F^{-1} H_\theta (w + \Delta u), \quad \theta \in \Theta, \quad w \in \mathbf{W}(\sigma), \quad \Delta \in \mathbf{\Delta}(\gamma) \tag{12}$$

where $\mathbf{W}(\sigma)$ is a set of bounded sequences with bounding parameter σ – not necessarily white noise sequences. Observe that now the model includes a dynamic uncertainty set, $\mathbf{\Delta}(\gamma)$, with γ denoting a system bound. The most common set used for robust control design is the set of $\mathbf{H}_\infty$ bounded systems, *i.e.*,

$$\mathbf{\Delta}(\gamma) = \left\{ \Delta(z) \;\middle|\; \|\Delta\|_\infty \le \gamma \right\} \tag{13}$$

with the corresponding prediction error,

$$e_\theta = F H_\theta^{-1} (y - G_\theta u) = w + \Delta u \tag{14}$$

In addition to the "random" component, $w \in \mathbf{W}(\sigma)$, the prediction error also has a "bias," or dynamic uncertainty component, $\Delta \in \mathbf{\Delta}(\gamma)$. In particular, with the true system given by (2), w and Δ are,

$$w = F H_\theta^{-1} v, \qquad \Delta = F H_\theta^{-1} (G - G_\theta) \tag{15}$$

In the ensuing discussion, let $\mathbf{W}(\sigma)$ be the set of sequences with bounded RMS, *i.e.*,

$$\mathbf{W}(\sigma) = \left\{ w \;\middle|\; \|w\|_{\rm rms} \le \sigma \right\}, \qquad \|w\|_{\rm rms}^2 := \lim_{N \to \infty} \frac{1}{N} \sum_{t=1}^{N} w(t)^2 \tag{16}$$

When the parametric model has the ARX structure, the uncertainty model becomes:

$$A_\theta y = B_\theta u + F^{-1}(w + \Delta u), \quad \|w\|_{\rm rms} \le \sigma, \quad \|\Delta\|_\infty \le \gamma \tag{17}$$

with $\theta \in \mathbf{R}^p$ and the corresponding prediction error,

$$e_\theta = F(A_\theta y - B_\theta u) = w + \Delta u \tag{18}$$

If the true system is given by (2), then w and Δ become,

$$w = F A_\theta v, \qquad \Delta = F(A_\theta G - B_\theta) \tag{19}$$

Observe that if F is a lowpass filter, then outside the filter bandwidth the uncertainty can be very large, whereas inside the filter bandwidth the uncertainty could be modest.

Unfalsification

Given some input/output data, the problem is to find a consistent set of uncertainty model parameters (θ, γ, σ). This problem is similar to the *uncertainty model validation problem* addressed by Poolla *et al.* [13]. In the case investigated there, the uncertainty model parameters are known and the uncertainty model is validated (unfalsified) from data. We use the same theory here (in reverse) to estimate the parameters.

We first define the unfalsification problem for the PE uncertainty model (12), and then specialize to the ARX case.

- **PE Uncertainty Model Unfalsification**

The uncertainty model (12)

$$y = G_\theta u + F^{-1} H_\theta (w + \Delta u), \quad \theta \in \Theta, \quad \|w\|_{\rm rms} \le \sigma, \quad \|\Delta\|_\infty \le \gamma$$

is said to be *unfalsified* by the N-point data sequences (1),

$$y^N = \{y(t)\}_{t=1}^N, \qquad u^N = \{u(t)\}_{t=1}^N$$

if and only if there exists $\theta \in \Theta$, an N-point sequence $w^N = \{w(t)\}_{t=1}^N$, and a system Δ, such that:

$$\|\Delta\|_\infty \le \gamma \quad \text{with} \quad \Delta : u^N \mapsto e_\theta^N - w^N \tag{20}$$

and

$$\left\| w^N \right\|_{\rm rms} \le \sigma \tag{21}$$

where e_θ^N is the N-point prediction error data sequence,

$$e_\theta^N = F H_\theta^{-1} (y^N - G_\theta u^N) \tag{22}$$

The key for establishing whether an uncertainty model is unfalsified is to prove (20). For LTI systems we have the following result which is stated in [13].

- **LTI Unfalsification** Given N-point sequences $u^N = \{u(t)\}_{t=1}^N$ and $z^N = \{z(t)\}_{t=1}^N$, there exists a stable, causal, LTI system such that

$$\|\Delta\|_\infty \le \gamma \quad \text{and} \quad z^N = \Delta x^N$$

if and only if (23)

$$\mathcal{T}(z^N)^T \mathcal{T}(z^N) \le \gamma^2 \mathcal{T}(u^N)^T \mathcal{T}(u^N)$$

where $\mathcal{T}(\cdot)$ maps an N-point sequence into an $N \times N$ lower triangular Toeplitz matrix, *e.g.*,

$$\mathcal{T}(u^N) = \begin{bmatrix} u(1) & 0 & \cdots & 0 \\ u(2) & u(1) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ u(N) & u(N-1) & \cdots & u(1) \end{bmatrix} \tag{24}$$

Applying the above result yields the following:

- **PE Uncertainty Model Unfalsification** The PE uncertainty model

$$y = G_\theta u + F^{-1} H_\theta (w + \Delta u), \quad \theta \in \Theta, \quad \|w\|_{\rm rms} \le \sigma, \quad \|\Delta\|_\infty \le \gamma$$

is unfalsified by the N-point data sequences (y^N, u^N) if and only if there exists $\theta \in \Theta$ and an N-point sequence $w^N = \{w(1), \ldots, w(N)\}$ such that:

$$\begin{aligned} \mathcal{T}(e_\theta^N - w^N)^T \mathcal{T}(e_\theta^N - w^N) &\le \gamma^2 \mathcal{T}(u^N)^T \mathcal{T}(u^N) \\ \frac{1}{N} \mathrm{vec}(w^N)^T \mathrm{vec}(w^N) &\le \sigma^2 \end{aligned} \tag{25}$$

From (23), the first inequality (25) is necessary and sufficient for the existence of an uncertain dynamic system $\|\Delta\|_\infty \le \gamma$ which maps $u^N \to e_\theta^N - w^N$. The second inequality in (25) is (obviously) necessary and sufficient for the existence of an N-point disturbance sequence satisfying $\left\| w^N \right\|_{\rm rms} \le \sigma$.

As shown in the recent text by Boyd *et al.* [2], the inequalities in (25) can also be expressed as matrix inequalities, *e.g.*,

$$\begin{bmatrix} \gamma\, \mathcal{T}(u^N)^T\mathcal{T}(u^N) & \mathcal{T}(e_\theta^N - w^N)^T \\ \mathcal{T}(e_\theta^N - w^N) & \gamma\, I_N \end{bmatrix} \geq 0$$

$$\begin{bmatrix} \sigma & \frac{1}{\sqrt{N}}\mathrm{vec}(w^N)^T \\ \frac{1}{\sqrt{N}}\mathrm{vec}(w^N) & \sigma\, I_N \end{bmatrix} \geq 0 \qquad (26)$$

For the general PE model structure, these matrix inequalities are affine in the unknown variables $\gamma \in \mathbf{R}_+$, $\sigma \in \mathbf{R}_+$, and $\mathrm{vec}(w^N) \in \mathbf{R}^N$, but *not* in the parameter $\theta \in \mathbf{R}^p$, except in the ARX case which we will discuss in the next section.

Uncertainty Tradeoff Curve It is clear that for any given value of γ there is a minimum value of σ, the disturbance bound, which is consistent with the data. For each dynamic uncertainty bound γ, denote this value by $\sigma(\gamma)$. The graph of $\sigma(\gamma)$ versus γ establishes a tradeoff curve between model uncertainty, γ, and disturbance uncertainty σ. The curve also separates the unfalsified and falsified uncertainty models based on the current data. The curve $\sigma(\gamma)$ will be referred to as the *unceretainty tradeoff curve.*

The general PE uncertainty tradeoff curve is obtained by solving the following optimization problem:

- **Unfalsified PE uncertainty Tradeoff Curve**
  Fix γ and perform the optimization:

$$\sigma(\gamma) := \min_{\theta,\,\sigma,\,w^N} \sigma, \quad \text{subject to (26)} \qquad (27)$$

  Denote the minimizing values of θ and w^N by $\theta(\gamma)$ and $w^N(\gamma)$.

Since e_θ^N is not affine in θ for the general PE structure, it follows that (27) is not a convex optimization and the tradeoff function, $\sigma(\gamma)$, is therefore not necessarily convex in γ, but it is piece-wise continuous.

The range of γ values must also be determined. We will examine the specific properties of the tradeoff curve for the ARX uncertainty model.

ARX Uncertainty Model The above result applies to unfalsification of the ARX uncertainty model (17) with $\Theta = \mathbf{R}^p$. However, in the ARX case, e_θ^N is affine in θ. Hence, the matrix inequalities (26) become *linear* matrix inequalities (LMI's) in the unknown variables $\gamma \in \mathbf{R}_+$, $\sigma \in \mathbf{R}_+$, $\theta \in \mathbf{R}^p$, and $\mathrm{vec}(w^N) \in \mathbf{R}^N$. Consequently, to find those (θ, γ, σ) satisfying (25) requires solving a convex feasibility problem [2].

The ARX uncertainty tradeoff curve is then obtained by solving the following *convex* optimization problem:

- **Unfalsified ARX uncertainty Tradeoff Curve**
  Fix γ and perform the optimization:

$$\sigma(\gamma) := \min_{\theta,\,\sigma,\,w^N} \sigma, \quad \text{subject to the LMI's in (26)} \qquad (28)$$

  Denote the minimizing values of θ and w^N by $\theta(\gamma)$ and $\widehat{w}^N(\gamma)$.

Since (28) is a convex optimization, it follows that the tradeoff function, $\sigma(\gamma)$, is convex in γ. The model uncertainty bound, γ, can range between the two obvious extremes: when the prediction error is due only to the disturbance and when it is due only to the dynamic uncertainty.

The first extreme point is when all the prediction error is assumed to be due to the disturbance, *i.e.*, when $\gamma = 0$. It follows from (28), that $\gamma = 0$ if and only if $w^N = e_\theta^N$. Hence, we obtain the usual least-squares/ARX solution:

$$\begin{aligned} \sigma(0) &= \sigma_{\mathrm{ls}} := \min_\theta \left\| e_\theta^N \right\|_{\mathrm{rms}} \\ \theta(0) &= \theta_{\mathrm{ls}} := \arg\min_\theta \left\| e_\theta^N \right\|_{\mathrm{rms}} \end{aligned} \qquad (29)$$

Consequently, we can now provide yet another interpretation of least-squares: *it produces the minimum variance unfalsified uncertainty model.* Thus, no ARX model, $A_\theta y = B_\theta u + F^{-1}w$, can reproduce the data with a noise level $\|w\|_{\mathrm{rms}}$ smaller than σ_{ls}.

The second extreme point on the tradeoff curve is when all the prediction error is assumed to be due to the dynamic uncertainty, *i.e.*, when $\sigma = 0$. Let $\sigma(\bar{\gamma}) = 0$ where $\bar{\gamma}$ is obtained by solving (28) with $w^N = 0$. That is,

$$\bar{\gamma} = \min_{\theta,\,\gamma} \gamma, \quad \text{subject to} \quad \mathcal{T}(e_\theta^N)^T\mathcal{T}(e_\theta^N) \leq \gamma^2\mathcal{T}(u^N)^T\mathcal{T}(u^N) \qquad (30)$$

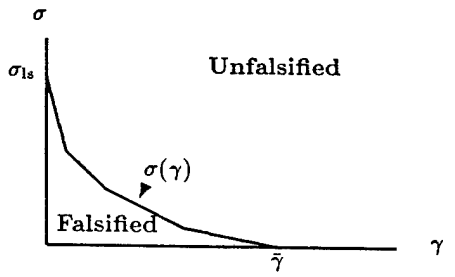
It is not always the case that a solution exists for $\sigma = 0$. As will be shown later, there may be a non-zero lower bound, $\bar{\sigma}$, such that $\sigma(\gamma) = \bar{\sigma}$ for $\gamma \geq \bar{\gamma}$. In any case, the unfalsified tradeoff curve is given more explicitly by:

$$\{\sigma(\gamma) \mid \gamma \in [0, \bar{\gamma}]\} \qquad (31)$$

A consequence of the convexity of $\sigma(\gamma)$, is that the two extreme points establish the affine upper bound:

$$\sigma(\gamma) \leq \bar{\sigma}(\gamma) := \sigma_{\mathrm{ls}}\,(1 - \gamma/\bar{\gamma}), \quad 0 \leq \gamma \leq \bar{\gamma} \qquad (32)$$

The figure below depicts a typical tradeoff curve separating unfalsified from falsified uncertainty models.



Frequency-Domain Analysis

The time-domain conditions for uncertainty model unfalsification can be transformed into the frequency domain by application of the discrete-Fourier-transform (DFT). This will also provide further insight into the meaning of the tradeoff curve as well as providing an efficient means for computation.

Discrete-Fourier-Tranform (DFT) Given an N-point time-domain sequence $x^N = \{x(t)\}_{t=1}^N$, define the function $X_N(\omega)$ by

$$X_N(\omega) := \frac{1}{\sqrt{N}} \sum_{t=1}^{N} x(t) e^{-j\omega t} \qquad (33)$$

The DFT of x^N is defined as the N-point complex sequence

$$\{X_N(\omega_n)\}_{n=1}^N \tag{34}$$

where the N-frequencies, $\omega_n = \frac{2\pi n}{N}$, $n = 1, \ldots, N$ are referred to as the DFT frequencies. The time sequence, $x^N = \{x(t)\}_{t=1}^N$, is completely recoverable from the DFT via the inverse transform:

$$x(t) = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} X_N(\omega_n) e^{j\omega_n t}, \quad t = 1, \ldots, N \tag{35}$$

Relation of DFT and Toeplitz Form The following result is established in [5]:

$$\lim_{N\to\infty} \lambda_{\max}\left(\frac{1}{N}\mathcal{T}(x^N)^T\mathcal{T}(x^N)\right) = \sup_{\omega} \lim_{N\to\infty} |X_N(\omega)|^2 \tag{36}$$

where $\lambda_{\max}$ denotes the maximum eigenvalue. Hence, for sufficiently large N,

$$\lambda_{\max}\left(\frac{1}{N}\mathcal{T}(x^N)^T\mathcal{T}(x^N)\right) \approx \max_n |X_N(\omega_n)|^2 \tag{37}$$

LTI Unfalsification The above approximation can be applied to LTI unfalsification (23). It is shown in [6] that for sufficiently large N, the data sequences u^N and z^N unfalsify the LTI uncertainty model $z^N = \Delta u^N$ with $\|\Delta\|_\infty \le \gamma$ if and only if,

$$\max_{n\in\mathcal{N}_U} \left|\frac{Z_N(\omega_n)}{U_N(\omega_n)}\right| \le \gamma, \quad \mathcal{N}_U := \{n \mid U_N(\omega_n) \neq 0\} \tag{38}$$

PE Tradeoff Curve Let $E_{\theta,N}(\omega_n)$, $U_N(\omega_n)$, $Y_N(\omega_n)$, and $W_N(\omega_n)$, denote the DFT's, respectively, of the N-point sequences e_θ^N, u^N, y^N, and w^N. The above approximation yield the following large-N approximation of (25):

$$\max_{n\in\mathcal{N}_U} \left|\frac{E_{\theta,N}(\omega_n) - W_N(\omega_n)}{U_N(\omega_n)}\right| \le \gamma, \quad \sum_{n=1}^{N} |W_N(\omega_n)|^2 \le \sigma^2 \tag{39}$$

Consequently, the large-N approximation of the PE tradeoff curve reduces to:

$$\sigma(\gamma)^2 \approx \min_\theta \left\{ \sum_{n\notin\mathcal{N}_U} |E_{\theta,N}(\omega_n)|^2 + \sum_{n\in\mathcal{N}_U} \left(|E_{\theta,N}(\omega_n)| - \min\left\{ \left|\frac{E_{\theta,N}(\omega_n)}{U_N(\omega_n)}\right|, \gamma \right\} |U_N(\omega_n)| \right)^2 \right\} \tag{40}$$

Of course no approximation is needed for computing $\sigma(0)$, which is precisely the least-squares solution, *i.e.*, by Parseval's Theorem,

$$\sigma(0)^2 = \sigma_{\rm ls}^2 = \sum_{n=1}^{N} |E_{\theta_{\rm ls},N}(\omega_n)|^2 \tag{41}$$

In addition, the large-N approximation of the value of γ for which $\sigma(\gamma) = 0$, $\bar{\gamma}$ from (30), is given by,

$$\bar{\gamma} \approx \min_\theta \max_{n\in\mathcal{N}_U} \left|\frac{E_{\theta,N}(\omega_n)}{U_N(\omega_n)}\right| \tag{42}$$

and the corresponding disturbance bound is not necessarily zero, *i.e.*,

$$\sigma(\bar{\gamma})^2 \approx \min_\theta \sum_{n\notin\mathcal{N}_U} |E_{\theta,N}(\omega_n)|^2 \tag{43}$$

ARX Tradeoff Curve The previous expressions are of course identical in form for the ARX uncertainty model, with the notable exception that all the indicated optimizations involve convex programming.

Computational Issues The DFT approximations reveal some of the computational issues. First, with the ARX uncertainty model structure, the DFT approximations allow very rapid calculations to find the optimum [6]. However, it is well known that dividing the DFT of an output sequence by the DFT of an input sequence is a very noisy estimate of the system frequency response, *e.g.*, [11, Ch.6]. But this is precisely what is suggested by the dynamic uncertainty constraint in (39). Hence, even though the calculations are rapid, they could give quite misleading results. There are several ways to reduce the poor quality of this estimate. It is standard practice to replace the DFT with spectral calculations, which typically involve time and frequency window selection. This brings us back to some of the earlier methods for caculating model error, *e.g.*, [7]. We will not pursue this important practical aspect any further here, but [12] provides the framework under which the spectral calculations would fit under the unfalsification umbrella.

Iterative Unfalsified Robust Control

In this section we provide a brief discussion on how to select an uncertainty model on the tradeoff curve which is suitable for control design. Because any uncertainty model corresponding to a point along the tradeoff curve is unfalsified by the current data record, it follows that a corresponding robust controller is also unfalsified by the current data record. Hence, a controller can only be *falsified* by a new data record. However, a new data record is obtained when the robust controller is implemented. This simple observation leads immediately to the following natural iterative unfalsification and robust control design procedure:

- **Unfalsification** Based on the current finite data record, solve (28) for the unfalsified tradeoff curve. This step in the iterative procedure replaces the identification step.
- **Robust Control Design** Corresponding to each feasible point, $(\gamma, \sigma(\gamma))$ on the tradeoff curve (31), design a robust controller, *i.e.*, a controller which attains the best closed-loop performance possible for the given uncertainty model parameters. Hence, a *family of unfalsified robust controllers* are obtained which are parametrized by γ, the model uncertainty bound, A large value of model uncertainty will yield a *cautious control*, and conversely, a small value of uncertainty will yield a more *aggressive control*. Since the family of robust controllers are based on the estimated tradeoff curve, there is no guaranty that any will perform as predicted. Hence, the controllers can only be falsified by real-time implementation.[1]
- **Real-Time Control Implementation** The uncertainty "knob," γ, allows the unfalsified controllers to be implemented in real-time in a logical order, *i.e.*, the first controller implemented should have a large uncertainty

[1] At this stage the methods of controller falsification proposed by Safonov *et al.* [14, 17, 18] could be used to falsify candidate controllers before they are implemented.

setting. If the actual performance is better than the predicted performance, then the controller (and the underlying uncertainty model) are still unfalsified. In this case, the controller can be made more aggressive by gradually decreasing the uncertainty level until either:

- a desired closed-loop performance is achieved.
- actual performance is worse than the predicted performance.

If the first condition occurs, controller design iterations cease and performance is still monitored until the second condition occurs, if ever. If the second condition occurs, the controller (and the underlying uncertainty model) have been *falsified*. In this case, the closed-loop data can be used to generate a new estimate of the unfalsified family of tradeoff curves. If these new curves are not significantly different than the previous curves, then more and more data can be used with perhaps more vigorous probing signals. If this fails, then the uncertainty model can be modified in a number of ways, *e.g.*, increase the nominal (ARX) model order, reshape or change the bandwidth of the data filter, modify input spectra, *etc.*

It is certainly conceivable that performance cannot be further improved because the limit of performance has been achieved. *But this cannot be known without knowing the system in the first place.* It is also possible that more fundamental changes must be made to the uncertainty model. For example, imposing an RMS bound on the disturbance uncertainty may be too restrictive. This can be relaxed by adding the constraint that the disturbance is uncorrelated with the input – as would be the case for stochastic signals. This is easily added to the tradeoff optimization problem via the convex constraint:

$$(w^N)^T u^N = 0$$

The disturbance can be completely altered by postulating a set of disturbances with bounded spectrum (see [12, 6]) and including the above constraint as well.

Stochastic sets are also possible, and again these would lead to convex optimization problems where now there would be a "probability tag" associated with the family of tradeoff curves and corresponding robust controllers.

Finally, failing all else, if the designer is "convinced" that the limit of performance has not been achieved, then it is possible that the underlying *linear* system assumption is wrong. That is, a linear model cannot capture some nonlinear phenomena that was presumed to be negligible, but is actually significant, and is hindering further performance improvement. In this case, there is strong motivation to re-think the physical modeling assumptions and to conduct experiments with small signal perturbations about nearby operating points to see if the tradeoff curves vary significantly. If so, this also suggests that there is an important nonlinear effect which should be accounted for directly in the control design, *e.g.*, gain scheduling, feedback linearization, *etc.*

References

[1] K. J. Åström and J. Nilsson, "Analysis of a scheme for iterated identification and control," *10th IFAC Symposium on System Identification*, vol. 2, pp. 171-176, Copenhagen, Denmark, 4-6 July 1994.

[2] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics, 1994.

[3] M.A. Dahleh and J.C. Doyle, "From data to control", *Proc. Workshop on Modeling of Uncertainty in Control Systems*, Springer-Verlag, 1992.

[4] M. Gevers, "Towards a joint design of identification and control," *Essays on Control: Perspectives in the Theory and its Applications*, editors: H.L. Trentelman and J.C. Willems, pp. 111-151, Birkhauser, Boston, MA, 1993.

[5] Ulf Grenander and Gabor Szego, *Toeplitz Forms and Their Applications*, Chelsea Pub. Co. 1st Edition 1958, 2nd Edition 1984.

[6] R.L. Kosut, "Iterative System Identification and Robust Control Design," in preparation.

[7] R.L. Kosut (1986), "Adaptive Calibration: An Approach to Uncertainty Modeling and On-Line Robust Control Design," *Proc. 25th IEEE CDC,* Athens, Greece, Dec. 1986.

[8] W.S. Lee, B.D.O. Anderson, and R.L. Kosut, "On adaptive robust control and control-relevant system identification", Tech. Report, Dept. of Systems Engr., Australian National University, 1992.

[9] M.M. Livestone, M.A. Dahleh, and J.A. Farrell, "A framework for robust control based model invalidation", *Proc. 1994 ACC*, pp. 3017-3020, Baltimore, MD, June 1994.

[10] M.M. Livestone, *Identification, Robust Adaptation, and Iterative Schemes*, Ph.D. dissertation, MIT Report LIDS-TH-2273, Oct. 1994.

[11] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, 1987.

[12] M. Massoumnia and R.L. Kosut, "A family of norms for system identification problems," *Proc. 1993 ACC*, San Francisco, CA, June 1993, and *IEEE Trans. Aut. Contr.*, vol. 39, no. 5, pp. 1027-1031, May, 1994.

[13] K. Poolla, P. Khargonnekar, A. Tikku, J. Krause, and K. Nagpal, "A time-domain approach to model validation", *Proc. 1992 ACC*, Chicago, IL, June 1992, and *IEEE Trans. Aut. Contr.*, vol. 39, no. 5, pp. 951-959, May 1994.

[14] M. G. Safonov and T. C. Tsao, "The unfalsified control concept: a direct path from experiment to controller", presented at *Conference on Feedback Control, Nonlinear Systems, and Complexity*, McGill University, Montreal,Canada, May 6-7, 1994.

[15] R.J.P. Schrama, "Accurate identification for control: the necessity for an iterative scheme", "Special Issue on System Identification for Robust Control Design," *IEEE Trans. Aut. Contr.*, vol. 37, no. 7, July, 1992.

[16] R.S. Smith and J.C. Doyle, "Model invalidation – a connection between robust control and identification", *Proc. 1989 ACC*, pp.1435-1440, 1989.

[17] T. C. Tsao and M. G. Safonov, "Set theoretic adaptor systems", *Proc. 1993 ACC*, pp. 3043-3047, San Francisco, CA, June 1993.

[18] T. C. Tsao and M. G. Safonov, "Convex set theoretic adaptor control systems", *Proc. 1993 CDC*, pp. 582-584, San Antonio, TX, Dec 1993.

[19] B. Wahlberg, "On model reduction in system identification," *Proc. 1986 ACC*, pp. 1260-1266, Seattle, WA, June 1986.

[20] Z. Zang, R.R. Bitmead, and M. Gevers, "Iterative model refinement and control robustness enhancement", Tech. Report, Dept. of Systems Engineering, ANU, Nov. 1991.

A.8 Iterative adaptive control via uncertainty model unfalsification

R. L. Kosut invited session, *1996 IFAC World Congress*, San Francisco, CA, June 1996.

Iterative Adaptive Robust Control via Uncertainty Model Unfalsification

Robert L. Kosut *
Integrated Systems, Inc.
3260 Jay St., Santa Clara, CA 95054
kosut@isi.com

Abstract - Unfalsification replaces system identification when the intended use of the model is robust control design. Unfalsification results in a family of uncertainty models, all consistent with the data, which tradeoff model uncertainty and disturbance uncertainty. The family of unfalsified models is used in an iterative approach to system identification and robust control design.

1 Introduction

> "*The test of all knowledge is experiment.* Experiment is the *sole judge* of scientific 'truth.' But what is the source of knowledge? Where do the laws that are to be tested come from? Experiment, itself, helps to produce these laws, in the sense that it gives us hints. But also needed is *imagination* to create from these hints the great generalizations – to guess ... and then to experiment to check again whether we have made the right guess."
>
> – Richard P. Feynman [3, Vol. I]

Applying Feynman's inspirational description of the Scientific Method to the problem of finding a model from data gives the algorithm: record data, pick a model, validate the model, if the model is no good pick a new model, validate the new model with new data, and so on. It is helpful of course, as Feynman puts it, to use one's "imagination" to arrive at "the great generalizations." For the problem addressed here, this means pick a model which is sufficiently accurate for use in control design. This paper is about how to do that.

The work described here is a continuation of [6] which addressed the problem of how uncertainty model unfalsification, as first described by Poolla *et al.* [13], could be used to replace the system identification step in adaptive control. In this paper the connection between robust control and unfalsification is made more precise. An iterative adaptive robust control algorithm is presented which incorporates the model falsification of Dahleh *et al.* [2, 9], the controller falsification approach of Safonov *et al.* [14], and some elements of the "windsurfer" adaptation devised by Anderson and Kosut as explained in [7, 8]. Due to limited space, the contributions of these and other references will be cited as needed in the text – a more complete listing and discussion can be found in [6], which is briefly reviewed in the next section.

*Research supported by the Air Force Office of Scientific Research, Directorate of Mathematical & Computer Sciences, under Contract No. F49620-93-C-0012.

2 Uncertainty Modeling

The starting point for robust control design is an *uncertainty model* which accounts for parametric, dynamic, and disturbance uncertainty. The specific form considered here (described in detail in [6]) is referred to as a *prediction error (PE) uncertainty model*:

$$y = G_\theta u + H_\theta(w + \Delta u) \quad \left\{ \begin{array}{c} \theta \in \Theta \\ \|w\|_{\rm rms} \le \sigma \\ \|\Delta\|_{\mathbf{H}_\infty} \le \delta \end{array} \right. \qquad (1)$$

where y and u are, respectively, the observed output and input sequences, w is an uncertain sequence rms-bounded by σ, Δ is an uncertain transfer function $\mathbf{H}_\infty$-bounded by δ, G_θ and H_θ are causal, linear-time-invariant systems, initially at rest, with transfer functions $G_\theta(z)$ and $H_\theta(z)$, each dependent on a parameter vector θ in the set,

$$\Theta = \left\{ \theta \in \mathbf{R}^p \;\middle|\; H_\theta^{-1} \text{ and } H_\theta^{-1} G_\theta \text{ are stable} \right\} \qquad (2)$$

Parameters in Θ insure that the *predictor* associated with (1) is stable [11], *i.e.*,

$$e_\theta := H_\theta^{-1}(y - G_\theta u) \qquad (3)$$

Observe that the PE uncertainty model (1) is characterized by *three* types of parameters, (θ, σ, δ) – the standard PE model is characterized by two types of parameters, (θ, σ).

3 Unfalsification

For robust control design, the parameters (θ, σ, δ) are known. In some cases θ can also be uncertain, *i.e.*, known

only to be in a subset of Θ. This is important particularly if θ represents uncertain physical parameters. In this paper we restrict attention to the problem of estimating (θ, σ, δ) from data. Hence, θ are used here as *canonical* parameters only to define a center for the uncertainty set.

Classical system identification is not the solution. It poses an *optimization problem* in (θ, δ), and it does not deal with the dynamic uncertainty $\|\Delta\|_{\mathbf{H}_\infty} \leq \delta$, which is of critical importance for robust control.

In contrast, a model is said to be *validated* if and only if it could have produced the data. Validation is perhaps a misnomer, as one can never *prove* that a model will be able to accurately predict the future. More precisely, and perhaps closer to Feynmans thoughts, the data can *falsify* a model, *i.e.*, the model may prove to be incapable of fully explaining the experiment. Hence, instead of validation, we use the more precise, but awkward term: *unfalsification.*

Clearly unfalsification is a *feasibility problem* – find a the model set whose members are consistent with the data. As shown in [6, 13], this philosophical shift allows the dynamic uncertainty bound δ to be estimated (unfalsified) along with θ and σ.

Specifically, let u^N and y^N denote, respectively, input and output data which have been recorded at N integer sample times $t = 1 : N$, that is,

$$y^N = \{y(t)\}_{t=1}^N, \quad u^N = \{u(t)\}_{t=1}^N \tag{4}$$

Let e_θ^N denote the corresponding prediction error sequence, *i.e.*,

$$e_\theta^N = H_\theta^{-1}(y^N - G_\theta u^N) \tag{5}$$

Unfalsification is defined as follows [6, 13].

- **Unfalsification**

  The uncertainty model (1) is said to be **unfalsified** by the N-point data sequences (y^N, u^N) if and only if there exists θ, $\Delta(z)$, and an N-point sequence w^N, such that: $\theta \in \Theta$, $\|\Delta\|_{\mathbf{H}_\infty} \leq \delta$, $\|w^N\|_{\text{rms}} \leq \sigma$, and (1) is consistent with the data, *i.e.*,

$$e_\theta^N = w^N + \Delta u^N \tag{6}$$

The following data dependent test of unfalsification is taken from [6], and is derived from the results in [13].

- **PE Uncertainty Model Unfalsification** The PE uncertainty model (1) is unfalsified by the N-point data sequences (y^N, u^N) if and only if there exists $\theta \in \Theta$ and an N-point sequence w^N such that:

$$\begin{aligned} \mathcal{T}(e_\theta^N - w^N)^T \mathcal{T}(e_\theta^N - w^N) &\leq \delta^2 \mathcal{T}(u^N)^T \mathcal{T}(u^N) \\ \frac{1}{N}\text{vec}(w^N)^T \text{vec}(w^N) &\leq \sigma^2 \end{aligned} \tag{7}$$

where $\mathcal{T}(\cdot)$ maps an N-point sequence into an $N \times N$ lower triangular Toeplitz matrix , *e.g.*,

$$\mathcal{T}(u^N) = \begin{bmatrix} u(1) & 0 & \cdots & 0 \\ u(2) & u(1) & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ u(N) & u(N-1) & \cdots & u(1) \end{bmatrix} \tag{8}$$

As shown in [6], an uncertainty tradeoff curve is obtained by solving the following optimization problem:

- **Unfalsified Tradeoff Curve**

  Fix δ and perform the optimization:

$$\sigma(\delta) := \min_{\theta, \sigma, w^N} \sigma, \quad \text{subject to (7)} \tag{9}$$

  Denote the minimizing values of θ and w^N by $\theta(\delta)$ and $w^N(\delta)$.

Remarks

(1) The graph of $\sigma(\delta)$ versus δ, referred to as the *uncertainty tradeoff curve*, establishes the tradeoff between model uncertainty, δ, and disturbance uncertainty σ. To every point on the tradeoff curve there is a different set of nominal transfer functions (G_θ, H_θ) because $\theta = \theta(\delta)$.

(2) The tradeoff curve separates the unfalsified and falsified uncertainty models based on the current data. (Figure 1 shows a typical tradeoff curve for the ARX uncertainty model). Solutions in the unfalsified (feasibility) set may not be unique. In addition, there is no easy parametrization, and attempts to discretize the space may lead to huge dimensions even for a few parameters.

(3) Let (θ, σ, δ) denote the parameters of the uncertainty model (1) satisfying (7). This is essentially a feasibility set because all that is required is the existence of (Δ, w, θ) such that $e_\theta^N = w + \Delta u^N$, $\sigma(\delta) \leq \|w\|_{\text{rms}} \leq \sigma$ and $\|\Delta\|_{\mathbf{H}_\infty} \leq \delta$.

(4) The model uncertainty bound, δ, can range between the two extremes: when the prediction error is due only to the disturbance ($\delta = 0$) and when it is due only to the dynamic uncertainty ($\sigma = 0$). In the latter case,

$$\begin{aligned} \delta_0 \quad = \quad & \min_\theta \delta \\ & \text{subject to} \\ & \mathcal{T}(e_\theta^N)^T \mathcal{T}(e_\theta^N) \leq \delta^2 \mathcal{T}(u^N)^T \mathcal{T}(u^N) \end{aligned} \tag{10}$$

Unlike the whole of the unfalsified region, uncertainty models along the tradeoff curve are easily parametrized, *i.e.*, $\theta(\delta), \sigma(\delta) 0 \leq \delta\gamma_0$.

(5) For the special case when $\delta = 0$, the corresponding uncertainty model on the tradeoff curve defined by the parameter triple $(\theta(0), \sigma(0), 0)$ is precisely the usual least-squares prediction error transfer function estimate.

(6) In general, the tradeoff curve is not convex because e_θ is not affine in θ except in the ARX case, which is described below.

(7) Notice that no mention has been made about the "true" system which generated the data – it need not be linear. All that is claimed is that there exists a set of *linear uncertainty models*, each of which *could* have produced the data.

ARX Uncertainty Model The ARX model is the most widely used model structure for standard system identification [11]. In this case:

$$\begin{aligned} G_\theta(z) &= \frac{B_\theta(z)}{A_\theta(z)}, \quad H_\theta(z) = \frac{1}{F(z)A_\theta(z)} \\ A_\theta(z) &= 1 + a_1 z^{-1} + \cdots + a_n z^{-n_A} \\ B_\theta(z) &= b_1 z^{-1} + \cdots + b_n z^{-n_B} \\ \theta &= [a_1 \cdots a_{n_A}\ b_1 \cdots b_{n_B}]^T \in \mathbf{R}^p, \quad p = n_A + n_B \end{aligned}$$

where $F(z)$ is the *data filter* included as part of the noise model $H_\theta(z)$. Hence, the (filtered) ARX prediction error,

$$e_\theta = F(A_\theta y - B_\theta u) \tag{11}$$

is affine in θ, from which it follows that $\Theta = \mathbf{R}^p$, the tradeoff curve $\sigma(\delta)$ is convex in δ, and the optimizing values $\theta(\delta)$ and $w^N(\delta)$ are unique. Of practical importance is that the optimization problem (9) is now convex because the matrix inequalities (7) become *linear* matrix inequalities (LMI's) in the unknown variables $\delta \in \mathbf{R}_+$, $\sigma \in \mathbf{R}_+$, $\theta \in \mathbf{R}^p$, and $\mathrm{vec}(w^N) \in \mathbf{R}^N$. As shown in [1], (7) can also be expressed as

$$\begin{bmatrix} \delta\, \mathcal{T}(u^N)^T \mathcal{T}(u^N) & \mathcal{T}(e_\theta^N - w^N)^T \\ \mathcal{T}(e_\theta^N - w^N) & \delta\, I_N \end{bmatrix} \geq 0$$

$$\begin{bmatrix} \sigma & \frac{1}{\sqrt{N}}\mathrm{vec}(w^N)^T \\ \frac{1}{\sqrt{N}}\mathrm{vec}(w^N) & \sigma\, I_N \end{bmatrix} \geq 0 \tag{12}$$

This expression more clearly exposes the LMI structure when e_θ is affine in θ.

ARX Tradeoff Curve Figure 1 depicts a typical ARX uncertainty model tradeoff curve separating unfalsified from falsified uncertainty models.

Additional Constraints Imposing only an rms bound on the disturbance uncertainty may be too harsh, resulting in unrealistic (malicious) disturbances. A more realistic situation is obtained by restricting the disturbance to be a zero-mean sequence independent of the input – as

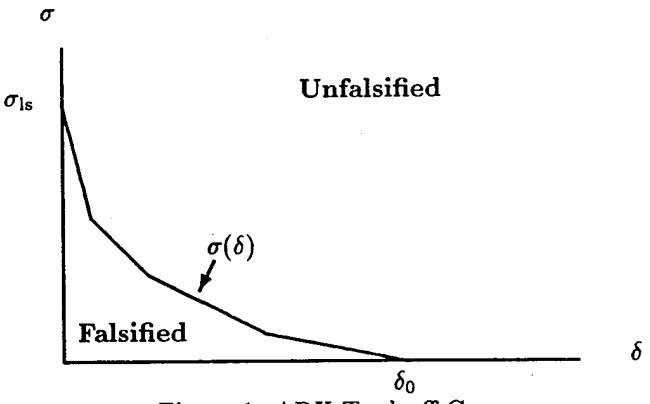


Figure 1: ARX Tradeoff Curve

would be the case for stochastic signals. These are easily incorporated in computing the tradeoff (9) by adding the *linear* constraints:

$$\textbf{zero-mean:} \quad \frac{1}{N}\sum_{t=1}^{N} w(t) = 0 \tag{13}$$

$$\textbf{independence:} \quad \frac{1}{N}\sum_{t=1}^{N} w(t)u(t) = 0 \tag{14}$$

Alternately, the disturbance can be completely characterized by bounding it's spectrum (see [12]).

Stochastic dynamic uncertainty descriptions as in [5] would lead to convex optimization problems where now there would be a "probability tag" associated with the family of tradeoff curves and corresponding robust controllers.

4 Unfalsified Robust Control

A robust control can be designed for every unfalsified uncertainty model and then a new closed-loop experiment can be run "to check again ..." But, as already remarked, the set of unfalsified uncertainty models satisfying (7) is too large for practical purposes. Therefore, candidate uncertainty models for robust control design will be chosen from among those along the tradeoff curve, *i.e.*, (1) with the parameter set $\{\theta(\delta), \sigma(\delta) \mid 0 \leq \delta \leq \delta_0\}$ found from (9). To emphasize that the candidate uncertainty models are unfalsified, as well as simplifying notation, let

$$G_{\mathrm{unf}} := G_{\theta(\delta)}, \quad H_{\mathrm{unf}} := H_{\theta(\delta)} \tag{15}$$

Let r denote the reference command and let,

$$u = K_{\mathrm{unf}}(r - y) \tag{16}$$

be an *unfalsified robust controller*, that is, K_{unf} stabilizes the candidate unfalsified uncertainty model. Hence, the closed-loop transfer functions,

$$T_{\mathrm{unf}} := \frac{G_{\mathrm{unf}}K_{\mathrm{unf}}}{1 + G_{\mathrm{unf}}K_{\mathrm{unf}}}, \quad Q_{\mathrm{unf}} := \frac{K_{\mathrm{unf}}}{1 + G_{\mathrm{unf}}K_{\mathrm{unf}}}, \tag{17}$$

$S_{\text{unf}} := 1 - T_{\text{unf}}$, $S_{\text{unf}} H_{\text{unf}}$, $S_{\text{unf}} G_{\text{unf}}$, and $Q_{\text{unf}} H_{\text{unf}}$ are all stable. In addition:

$$\delta \left\| Q_{\text{unf}} H_{\text{unf}} \right\|_{\mathbf{H}_\infty} < 1 \tag{18}$$

Recall that this condition is necessary and sufficient to insure stability of the closed-loop unfalsified uncertainty model for all $\|\Delta\|_{\mathbf{H}_\infty} \leq \delta$. However, it is neither necessary nor sufficient for determining if K_{unf} will *actually* stabilize the true system. Recall Feynman's words: "Experiment is the *sole judge* of scientific 'truth.' " So until we perform an experiment with K_{unf} as the controller, we cannot know if it is stabilizing even if (18) is false. However, there is a sensible order in which to experiment with candidate controllers, and as shown by Safonov *et al.* [14], some of them can be eliminated (falsified) without running a new experiment. We will return to this in the next section.

If there was no uncertainty in the model ($w = 0, \Delta = 0$), then the *nominal unfalsified* closed-loop response would be $T_{\text{unf}} r$. Non-zero disturbance and dynamic uncertainty yields the *unfalsified closed-loop performance error*:

$$\begin{aligned} e_{\text{unf}} &:= y - T_{\text{unf}} r \qquad (19) \\ &= \left(\frac{S_{\text{unf}} Q_{\text{unf}} H_{\text{unf}} \Delta}{1 + Q_{\text{unf}} H_{\text{unf}} \Delta} \right) r + \left(\frac{S_{\text{unf}} H_{\text{unf}}}{1 + Q_{\text{unf}} H_{\text{unf}} \Delta} \right) w \end{aligned}$$

The above expression can be thought of as an unfalsified uncertainty model of the closed-loop system performance error.

Under the assumptions that uncertainty (Δ, w) can be drawn from the sets $\|\Delta\|_{\mathbf{H}_\infty} \leq \delta$ and $\|w\|_{\text{rms}} \leq \sigma(\delta)$, and if (18) holds, then the unfalsified closed-loop system satisfies the following condition:
$\forall n \geq 1$ and $\forall \|r^n\|_{\text{rms}} < \infty$,

$$\|e^n_{\text{unf}}\|_{\text{rms}} \leq J_{\text{unf}}(T_{\text{unf}}, r^n) \tag{20}$$

where

$$\begin{aligned} J_{\text{unf}}(T, r) &:= \gamma_{\text{unf}} \|Tr\|_{\text{rms}} + \alpha_{\text{unf}} \\ \gamma_{\text{unf}} &:= \delta \left\| \frac{S_{\text{unf}} Q_{\text{unf}} H_{\text{unf}} / [T_{\text{unf}}]_m}{1 - \delta |Q_{\text{unf}} H_{\text{unf}}|} \right\|_{\mathbf{H}_\infty} \\ \alpha_{\text{unf}} &:= \sigma(\delta) \left\| \frac{S_{\text{unf}} H_{\text{unf}}}{1 - \delta |Q_{\text{unf}} H_{\text{unf}}|} \right\|_{\mathbf{H}_\infty} \end{aligned}$$

where $[T_{\text{unf}}]_m$ is the minimum phase factor of T_{unf}. The same caution regarding (18) applies here as well: there is no guaranty that (20) holds when a closed-loop experiment is run with K_{unf}.

5 Uncertainty Model Falsification

As shown in [9], if (20) does not hold with data from a new closed-loop experiment with K_{unf}, then the uncertainty model associated with K_{unf} is *falsified* by the new data. More precisely:

- **Model Falsification**

  Perform a closed-loop experiment with control K_{unf} and record new data $(y^N_{\text{expt}}, u^N_{\text{expt}}, r^N_{\text{expt}})$, and

  $$e^N_{\text{expt}} = y^N_{\text{expt}} - T_{\text{unf}} r^N_{\text{expt}} \tag{21}$$

  If $\exists n \in [1 : N]$ such that:

  $$\left\| e^n_{\text{expt}} \right\|_{\text{rms}} > J_{\text{unf}}(T_{\text{unf}}, r^n_{\text{expt}}) \tag{22}$$

  then the uncertainty model for which K_{unf} has been designed is *falsified* by the new data.

Remarks

(1) A more informative expression signifying (22) is: *the controller is falsified by the new data with respect to the performance criterion (20).*

(2) The data used for unfalsifying the uncertainty model used in the design of K_{unf} is different than data taken from the experiment with K_{unf} in the loop. Hence, if (22) holds then the new data set, which has falsified the previously unfalsified model, can be used to compute a new tradeoff curve. This observation suggests an iterative procedure for adaptive control and will be discussed in a later section.

(3) It is possible that (22) holds, *i.e.*, the uncertainty model is falsified by the new data, but performance satisfies a desired specification. For example, suppose (22) holds and $\forall n \in [1 : N]$,

$$\left\| e^n_{\text{expt}} \right\|_{\text{rms}} \leq J_{\text{spec}}(T_{\text{unf}}, r^n_{\text{expt}}) \tag{23}$$

where

$$J_{\text{spec}}(T, r) := \gamma_{\text{spec}} \|Tr\|_{\text{rms}} + \alpha_{\text{spec}} \tag{24}$$

In this case the experiment provides a *proof* that the controller K_{unf} meets a desired specification, at least for one sequence r^N_{expt}. If this sequence is sufficiently rich in spectral content, it is highly unlikely that further experimentation is required to falsify K_{unf} with repect to $J_{\text{spec}}(T_{\text{unf}}, r)$. However, a different unfalsified model could still result in better performance – this has not been ruled out.

(4) The above experimental validation (falsification) of the controller K_{unf} is *indirect*, because the controller is designed from an unfalsified uncertainty model. This can lead to a possibly wide gap between unfalsified predicted performance and actual performance. It is possible to directly falsify controllers without having to try them out on the actual system. This is described next.

6 Controller Falsification

In a series of papers by Safonov *et al.* (see [14] and the references therein), it is shown how to *directly* falsify a candidate controller before it is implemented. This is a significant result, because time consuming and unnecessary experimentation is eliminated. The procedure for controller falsification is essentially the same as that for uncertainty model unfalsification, but applied to the controller. What is different is that the "model" now becomes the controller, and the "uncertainty" becomes the closed-loop specification.

Specifically, suppose we are given data (y^N, u^N) and want to know, *without additional experimentation*, if the candidate controller K_{unf} will *fail* to meet a desired closed-loop specification. A simplified version of the procedure described in [14] is as follows.

- **Controller Falsification**

  A candidate controller K_{unf}, designed for an uncertainty model unfalsified by the data (y^N, u^N), is *falsified with repect to performance* $J_{\text{spec}}(T_{\text{unf}}, r)$ if $\exists r^N_{\text{bad}}$ satisfying

$$u^N = K_{\text{unf}}(r^N_{\text{bad}} - y^N) \tag{25}$$

  and $\exists n \in [1 : N]$ such that,

$$\|e^n_{\text{bad}}\|_{\text{rms}} > J_{\text{spec}}(T_{\text{unf}}, r^n_{\text{bad}}) \tag{26}$$

  with

$$e^N_{\text{bad}} := y^N - T_{\text{unf}} r^N_{\text{bad}} \tag{27}$$

Remarks

Controller falsification is a feasibility problem in the variable r^N_{bad}. Solving for r^N_{bad} from (25) gives,

$$r^N_{\text{bad}} = y^N + K^{-1}_{\text{unf}} u^N \tag{28}$$

Using (17) gives,

$$e^N_{\text{bad}} = S_{\text{unf}} (y^N - G_{\text{unf}} u^N) \tag{29}$$

$$T_{\text{unf}} r^N_{\text{bad}} = y^N - S_{\text{unf}}(y^N - G_{\text{unf}} u^N) \tag{30}$$

Observe that $(S_{\text{unf}}, S_{\text{unf}} G_{\text{unf}})$ are stable by construction (17), hence, $e^N_{\text{bad}}, r^N_{\text{bad}}$ are bounded even though $\|r^N_{\text{bad}}\|_{\text{rms}}$ could be very large.

It can be shown that e_{bad} is precisely the performance error that would have been obtained had the controller

$$u = K_{\text{unf}}(r_{\text{bad}} - y) \tag{31}$$

been applied to the unfalsified uncertainty model for which K_{unf} is designed. As a result, using (20),

$$\|e_{\text{bad}}\|_{\text{rms}} \leq J_{\text{unf}}(T_{\text{unf}}, r_{\text{bad}}), \tag{32}$$

The practical consequence is this: falsification [with repect to $J_{\text{unf}}(T_{\text{unf}}, r)$] of the controller (26) implies falsification of the model (22). This reinforces the notion expressed by Safonov *et al.* [14] that *direct* controller falsification is a sharper test than the *indirect* approach using model falsification (22).

7 Iterative Adaptive Control

The preceding discussion of model and controller falsification suggests a "natural" iterative approach to adaptive control. The idea is to verify, ultimately by experiment, the performance of all robust controllers designed for the family of unfalsified uncertainty models along the δ-dependent tradeoff curve. The *unfalsified* controllers should be either falsified or implemented in decreasing order of δ, *i.e.*, the first controller implemented should be a cautious controller with a large uncertainty setting. . More precisely, the control can be iteratively adapted by adhering to the following steps:

1. **initialize performance specification**

$$J_{\text{spec}}(T, r) = \gamma_{\text{spec}} \|Tr\|_{\text{rms}} + \alpha_{\text{spec}}$$

2. **record new N-point batch of data**

$$(y^N, u^N, r^N) = (y^N, u^N, r^N)_{\text{batch}}$$

3. **compute tradeoff parameters**

   use (9) with (y^N, u^N) to obtain:

$$\{\theta(\delta), \sigma(\delta) \mid 0 \leq \delta \leq \delta_0\}$$

4. **start with cautious controller**

   set $\delta = \delta_0$

5. **robust control design**

   (a) design controller K_{unf}
   (b) compute $J_{\text{unf}}(T_{\text{unf}}, r)$ from (20)

6. **controller falsification**

   if (26) holds, *i.e.*,

$$\exists n, \ \|e^n_{\text{bad}}\|_{\text{rms}} > J_{\text{spec}}(T_{\text{unf}}, r^n_{\text{bad}})$$

   then GO TO Step 10

7. **new experiment & record data**

   (a) implement K_{unf}
   (b) record closed-loop data $y^N_{\text{expt}}, u^N_{\text{expt}}, r^N_{\text{expt}}$

8. **evaluate performance**

   if (26) holds, *i.e.*,

$$\exists n, \ \|e^n_{\text{expt}}\|_{\text{rms}} > J_{\text{spec}}(T_{\text{unf}}, r^n_{\text{expt}})$$

   then:
   (a) reset to last controller satisfying J_{spec}
   (b) replace (y^N, u^N, r^N) with $(y^N, u^N, r^N)_{\text{expt}}$, and GO TO 3.

9. **model falsification**

   if

   $$\exists n,\ \left\|e^n_{\text{expt}}\right\|_{\text{rms}} > J_{\text{unf}}(T_{\text{unf}}, r^n_{\text{expt}})$$

   then replace (y^N, u^N, r^N) with $(y^N, u^N, r^N)_{\text{expt}}$, and GO TO 3

   else, if $J_{\text{unf}} < J_{\text{spec}}$, then replace J_{spec} with J_{unf}.

10. **decrease δ**

    if $\delta = 0$, GO TO Step 2

    else, decrease δ and GO TO Step 5

Remarks

(1) The algorithm falsifies and/or implements controllers along a tradeoff curve, and when these are exhausted, a new batch of data is recorded and the process is repeated. Controllers are implemented only when they fail to be falsified by current data with respect to the best observed performance.

(2) In each new experiment the reference command r^N_{expt} can consist of a nominal sequence, such as a step or square-wave, plus a small random component which could be added to the test reference and/or directly to the control input. Just as in classical system identification, a sufficiently rich excitation is required to obtain useful estimates, *e.g.*, [11, 4].

(3) For each unfalsified model the controller can be designed in any number of ways, and moreover, there are many possible robust controllers. For example, the problem formulation suggests a mixed $\mathbf{H}_2/\mathbf{H}_\infty$-optimization, but that involves choices of weightings, all of which effect T_{unf}, the nominal response to a reference. A more direct approach is suggested in [7, 8] where

$$Q_{\text{unf}} = T_\lambda / [G_{\text{unf}}]_m \tag{33}$$

where $[G_{\text{unf}}]_m$ denotes the minimum-phase factor of G_{unf}, and T_λ is the desired (minimum phase) closed-loop transfer function from r to y with bandwidth λ. The convenience of this approach is that the two important design "knobs" are directly accessible, namely, model error δ and desired closed-loop bandwidth λ.

(4) A large value of model uncertainty δ together with a low bandwidth λ will yield a *cautious control.* Conversely, a small value of uncertainty δ together with a large bandwidth λ will yield a more *aggressive control.* If, in addition, T_λ is used as the data filter when producing the tradeoff models, then it becomes easier to link unfalsified models with the desired performance.This additional degree of design freedom is easily added to the iterative algorithm.

Concluding Remarks

An iterative approach to robust adaptive control has been proposed where unfalsification of an uncertainty model replaces classical system identification. For the ARX model, the unfalsification step requires solving a set of convex programming problems, specifically LMI problems, of which ordinary least-squares is one member. The result is a tradeoff curve between model uncertainty and disturbance uncertainty. Hence, a family of models are unfalsified from the data record. A natural order of design is to start with a cautious controller designed for an unfalsified uncertainty model with a large value of dynamic uncertainty. As the model uncertainty is slowly decreased, actual closed-loop data is used to evaluate (unfalsify or falsify) the model and/or the controller.

References

[1] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, SIAM Studies in Applied Mathematics, 1994.

[2] M.A. Dahleh and J.C. Doyle, "From data to control", *Proc. Workshop on Modeling of Uncertainty in Control Systems*, Springer-Verlag, 1992.

[3] R. P. Feynman, R.B. Leighton, and M. Sands *The Feynman Lectures on Physics*, Addison-Wesley, 1963.

[4] M. Gevers, "Towards a joint design of identification and control," *Essays on Control: Perspectives in the Theory and its Applications*, editors: H.L. Trentelman and J.C. Willems, pp. 111-151, Birkhauser, Boston, MA, 1993.

[5] G.C. Goodwin, M. Gevers, and B. Ninness, "Quantifying the error in estimated transfer functions with applications to model order selection," *IEEE Trans. on Automatic Control*, vol. 37, no.7, pp. 913-928, July 1992.

[6] R.L. Kosut, "Uncertainty model unfalsification: a system identification paradigm compatible with robust control design," *Proc. 1995 CDC*, New Orleans, LA, Dec. 1995.

[7] W.S. Lee, B.D.O. Anderson, R.L. Kosut, and I.M.Y. Mareels, "A new approach to adaptive robust control," *Int. J. of Adaptive Control and Signal Processing*, vol. 7, pp. 183-211, 1993.

[8] W.S. Lee, B.D.O. Anderson, I.M.Y. Mareels, and R.L. Kosut, "On some key issues in the windsurfer approach to robust adaptive control", *Automatica*, Vol. 31, No. 11, pp. 1619-1636, 1995.

[9] M.M. Livestone, M.A. Dahleh, and J.A. Farrell, "A framework for robust control based model invalidation", *Proc. 1994 ACC*, pp. 3017-3020, Baltimore, MD, June 1994.

[10] M.M. Livestone, *Identification, Robust Adaptation, and Iterative Schemes*, Ph.D. dissertation, MIT Report LIDS-TH-2273, Oct. 1994.

[11] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, 1987.

[12] M. Massoumnia and R.L. Kosut, "A family of norms for system identification problems," *Proc. 1993 ACC*, San Francisco, CA, June 1993, and *IEEE Trans. Aut. Contr.*, vol. 39, no. 5, pp. 1027-1031, May, 1994.

[13] K. Poolla, P. Khargonnekar, A. Tikku, J. Krause, and K. Nagpal, "A time-domain approach to model validation", *Proc. 1992 ACC*, Chicago, IL, June 1992, and *IEEE Trans. Aut. Contr.*, vol. 39, no. 5, pp. 951-959, May 1994.

[14] M. G. Safonov and T. C. Tsao, "The unfalsified control concept: a direct path from experiment to controller", presented at *Conference on Feedback Control, Nonlinear Systems, and Complexity*, McGill University, Montreal,Canada, May 6-7, 1994.